DOCUMENT RESUME

ED 100 356 16 IR 001 454

AUTHOR Nickerson, Raymond S.; And Others
TITLE Draft Curriculum Manual: A Review of Speech Problems of the Deaf and a Preliminary Guide to the Use of the BBN System of Speech-Training Aids. Report No. 2911.
INSTITUTION Bolt, Beranek and Newman, Inc., Cambridge, Mass.
SPONS AGENCY Bureau of Education for the Handicapped (DHEW/OE), Washington, D.C. Media Services and Captioned Films Branch.
REPORT NO BBN-R-2911
PUB DATE 14 Sep 74
CONTRACT OEC-0-71-4670(615)
NOTE 114p.; For related documents see IR 001 449-455

EDRS PRICE MF-$0.75 HC-$5.40 PLUS POSTAGE
DESCRIPTORS *Computer Assisted Instruction; *Deaf Children; Deaf Education; Deaf Research; Educational Diagnosis; Handicap Detection; Literature Reviews; Material Development; *Speech Education; Speech Evaluation; *Speech Handicaps; Speech Instruction; Speech Skills; Teacher Developed Materials; *Teaching Guides

ABSTRACT

Material for teacher-users of a computer-based system of speech training aids for the deaf is offered. Research on the types of deficiencies found in the speech of the deaf is reviewed; a philosophy concerning training which emphasizes the role of diagnosis, is presented; and suggestions are made concerning use of the displays produced by the system to facilitate diagnosis and training. (Author/SK)

Report No. 2911

Bolt Beranek and Newman Inc.

ED100356

DRAFT CURRICULUM MANUAL: A REVIEW OF SPEECH PROBLEMS OF THE DEAF AND A PRELIMINARY GUIDE TO THE USE OF THE BBN SYSTEM OF SPEECH-TRAINING AIDS

R. S. Nickerson
K. N. Stevens
A. Boothroyd
R. E. Adams
R. D. Storm

14 September 1974

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

IR 001 454

Submitted to

Media Services and Captioned Films Branch
Division of Educational Services
Bureau of Education for the Handicapped
U.S. Department of Health, Education, and Welfare
Washington, D.C. 20202

Table of Contents

	Page
Abstract	iii
1. INTRODUCTION	1
2. PROBLEMS WITH SPEECH OF THE DEAF	3
2.1 Timing and Rhythm	3
2.2 Pitch and Intonation	7
2.3 Velar Control	14
2.4 Articulation	17
2.5 Voice Quality	23
2.6 Interrelatedness of Problems	26
3. TRAINING PHILOSOPHY	29
3.1 Diagnosis	29
3.2 Training Objectives	32
3.3 Training Procedures	34
3.4 Evaluation of Progress	35
4. USE OF DISPLAYS FOR SPEECH DIAGNOSIS AND TRAINING	36
4.1 Timing and Rhythm	37
4.2 Control of Fundamental Frequency (F_0)	52
4.3 Velar Control	65
4.4 Articulation	78
4.4.1 Fricative Consonants	83
4.4.2 Vowels and Diphthongs	90
4.5 Voice Quality	98
5. ACKNOWLEDGMENTS	101
6. REFERENCES	102

Abstract

Our primary purpose in writing this document was to provide some material that would be helpful to teacher-users of the BBN computer-based system of speech-training aids for the deaf. Research on the types of deficiencies that are found in the speech of the deaf is reviewed; a philosophy concerning training, and emphasizing the role of diagnosis, is presented; and suggestions are made concerning how the displays that can be produced by the system may be used to facilitate diagnosis and training. The tentativeness of the ideas that are presented in this document concerning training is acknowledged by the use of the word "draft" in the title.

1. INTRODUCTION

For over two years we have been attempting to develop, use and evaluate a computer-based system of speech-training aids for the deaf. The purpose of this document is to set forth what we think we have learned as a result of this effort concerning how such a system might be used to help teach speech to deaf children. Although the proposed training exercises and illustrations are all based on the capabilities of this particular system, it should not be inferred from this that we view this system as the ultimate in computer-based speech-training aids. Our feeling is quite to the contrary. We view this system as a very modest step in the direction of bringing computing technology to bear on the extraordinarily difficult problem that teaching speech to deaf children represents.

Our experience to date has convinced us of the validity of the assumption that computer technology has the potential to impact favorably and significantly on this problem, but it is also clear that a long time and considerable effort may be required to exploit that potential fully. The major difficulties are not technological limitations, but pedagogical uncertainties. Even with a relatively small computing machine such as the one in our system, we are able to generate more displays than we know how to use effectively. We take this not as cause for discouragement but as support for a point of view that was expressed at the beginning of this project; namely, that a system of the sort toward which we are working cannot be designed. Such a system must be evolved, and the evolutionary process must be guided by a close interaction between developers and users. This is the approach that has been taken in the development of the system with which this manual is concerned. We consider the system

to be still in an early stage of evolution, but it has evolved to a point at which it can, we think, be used to advantage in the day-to-day training of speech.

2. PROBLEMS WITH SPEECH OF THE DEAF

It may be somewhat misleading to talk about specific problems, or specific deficiencies, in the speech of the deaf, because it can create the impression that these problems are independent and can be dealt with individually. Independence is almost certainly not the case. No matter how one categorizes the various deficiencies that have been focused upon, one discovers that one cannot say much about any given category without implicating others. Nevertheless, some structuring is necessary.

In this report the discussion of speech problems is organized around five topics: timing and rhythm, pitch and intonation, velar control, articulation, and voice quality. This selection reflects nothing more than the fact that we find it convenient to structure our own thinking in this way. No claim is made for the superiority of this organization over others that might be used.

2.1 Timing and Rhythm

Poor timing has been considered by several investigators to be a major cause of the generally poor intelligibility of the speech of the deaf (Bell, 1916; Hood, 1966; Hudgins & Numbers, 1942; John & Howarth, 1965; Houde, 1973; Nober, 1967; Stratton, 1973).

Precise specification of timing deficiencies is not possible simply because not enough is yet known about the temporal characteristics of "normal" speech. The results from a number of studies permit several tentative assertions, however, that are at least suggestive of ways in which the speech of the deaf

may differ, in the aggregate, from that of hearing speakers with respect to temporal aspects. Many of these findings have been reviewed in greater detail by Nickerson, Stevens, Boothroyd, and Rollins (1974).

Deaf speakers tend to speak at a much slower rate than do hearing speakers (Boone, 1966; Colton & Cooker, 1968; Hood, 1966; John & Howarth, 1965; Martony, 1966; Mason & Bright, 1937; Nickerson, et al., 1974; Voelker, 1938). It has been estimated that hearing speakers emit speech at the rate of about 3.3 syllables (Pickett, 1968) or 10 to 12 phonemes (Miller, 1962) per second on the average; although speech rate can deviate considerably in either direction from these norms and still be highly intelligible (Abrams, Goffard, Kryter, Miller, Sanford, & Sanford, 1944). Deaf speakers tend to speak more slowly than the slowest hearing speakers, however, and when the deaf and hearing speakers have been studied under similar conditions, the measured rates have often differed by a factor of two or three, or more (Hood, 1966; Mason & Bright, 1937; Voelker, 1938).

Deaf speakers fail to make the difference between the durations of stressed and unstressed syllables sufficiently large (Angelocci, 1962; Nickerson, et al., 1974). This is because, although deaf speakers prolong the durations of both stressed and unstressed syllables relative to the hearing, the increase tends to be proportionally greater for unstressed than for stressed sounds. Hearing speakers lengthen stressed syllables and syllables in word-final and sentence-final positions (Parmenter & Trevino, 1935; Fry, 1958; Lindblom & Rapp, 1973; Klatt, 1974). A stressed syllable in final position is likely to be three to five times as long as a preceding

unstressed syllable; for deaf speakers the ratio is typically much smaller than this. It is almost as though the deaf speaker produces only stressed syllables; and, in fact, some investigators have suggested that this problem is in part a result of training that puts great emphasis on the articulation of individual speech sounds in isolation or in isolated consonant-vowel syllables (Boone, 1966; John & Howarth, 1965).

Deaf speakers tend to insert more pauses, and pauses of longer duration, in running speech than do hearing speakers (Hood, 1966; Hudgins, 1946; John & Howarth, 1965; Nickerson, et al., 1974). Moreover, these pauses often are inserted at inappropriate places, such as within phrases (Nickerson, et al., 1974).

Closely related to the problem of excessive and inappropriately placed pauses is that of poor rhythm. When listeners are asked to rate the adequacy of the rhythm or syllable grouping of the speech of deaf speakers, the ratings are below those for hearing speakers (Hood, 1966; Hudgins, 1946). The importance of speech rhythm for intelligibility has been demonstrated by Hudgins and Numbers (1942).

Both the problem of pauses and that of poor rhythm are related to, or perhaps result at least in part from, inadequate breath control during speech production (DiCarlo, 1964; Hudgins, 1934, 1936, 1937, 1946; Scuri, 1935; Rawlings, 1935, 1936). Apparently, deaf speakers expel much more breath while speaking than do hearing speakers (Hudgins, 1934, 1937, 1946; Rawlings, 1935, 1936), and consequently are likely to interrupt the speech flow more frequently in order to permit the intake of air.

Apparently, too, deaf speakers tend to use more breath during speaking than when not speaking; whereas hearing speakers use about the same amount in both cases (Scuri, 1935). Scuri's data suggest that the deaf sometimes seem to lack the ability to close the glottis completely, which could help account not only for the excessive expenditure of breath as a factor in producing poor timing and rhythm, but also for the quality problem of "breathy" voice.

Some timing problems that have been noted by investigators of the speech of the deaf are associated with the production of certain types of speech sounds. Fricative consonants, for example, may have an inordinately long duration for deaf speakers (Angelocci, 1962; Calvert, 1961, 1962), as may the closure periods of plosive consonants (Angelocci, 1962; Calvert, 1962).

Finally, speech sounds that require the precise coordination of the timing of different articulatory movements or the rapid transition from one articulatory position to another may be problematical for the deaf speaker. The timing of voice onset relative to release for a voiceless stop consonant (Angelocci, 1962), and that of the onset of nasalization for a nasal consonant (Stevens, Nickerson, Boothroyd, & Rollins, 1974) are cases in point, as is the timing of the movements required to produce consonant blends or diphthongs, or that of the transitions represented by the junctions between fricative or nasal consonants and vowels (Martony, 1966).

To teachers of speech to the deaf, timing problems can be particularly frustrating. Although timing deficiencies are widely considered to be especially important as causes of the lack of speech intelligibility, there is no well-developed

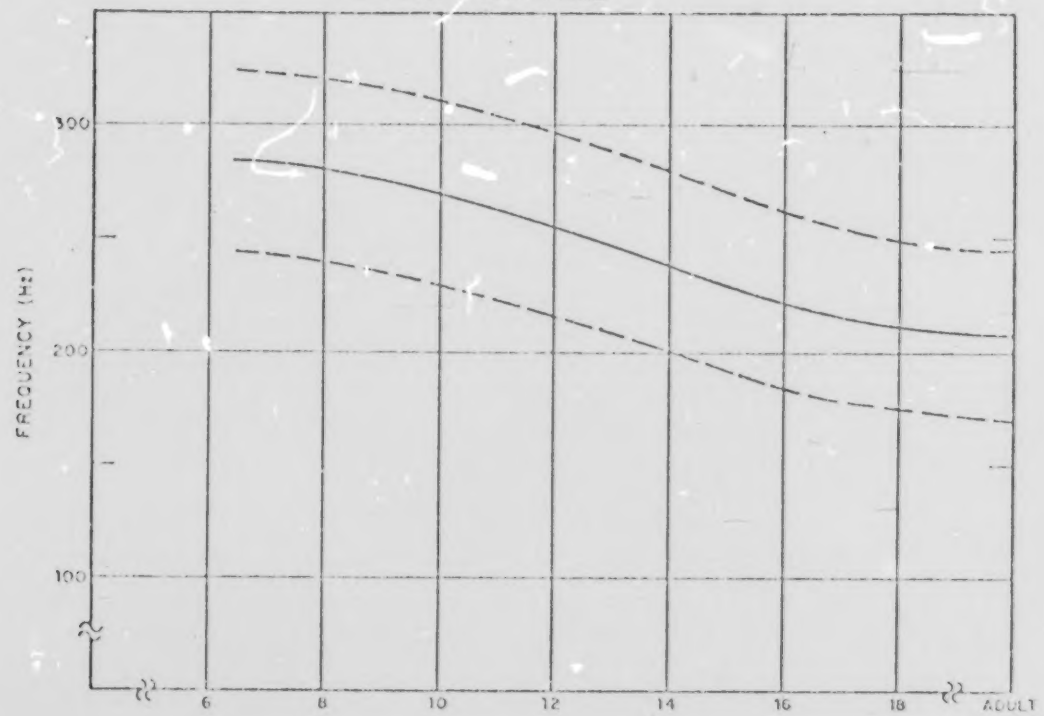
theory of timing from which performance objectives and evaluation criteria can be inferred. One may be able to say with confidence that the timing of a given utterance is grossly deficient, but it does not follow that one will be able to say with the same assurance precisely how it should be modified to make the timing right. Until an adequate theory of timing is developed, attempts to rectify timing deficiencies will, of necessity, be somewhat ad hoc.

Some progress toward a theory of speech timing is being made (Klatt, 1974; Lindblom & Rapp, 1973; see Nickerson, Stevens, Boothroyd, & Rollins, 1974, for a brief summary). It is probably important, however, to distinguish between a theory that is descriptive of the speech of the hearing and one that will provide the basis for establishing training goals for the deaf. It is not necessarily the case that normative timing patterns derived from statistical representations of the speech of hearing speakers constitute reasonable targets toward which all deaf children should be encouraged to strive. What is needed is a much better understanding of how intelligibility and quality of speech depend on various temporal features. Such understanding can only come from extensive studies of the speech of the deaf as well as of that of the hearing, designed to answer precisely this question.

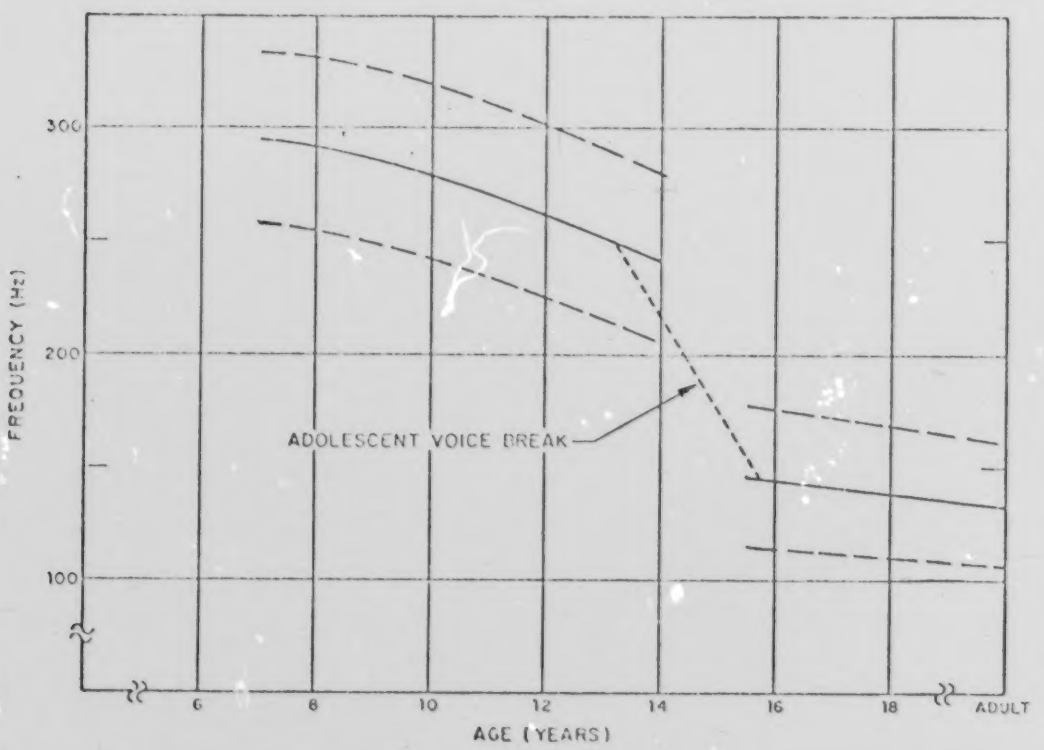
2.2 Pitch and Intonation

The fundamental frequency (F_0) of voiced speech sounds varies considerably in the speech of a given speaker, and the average, or characteristic, F_0 varies over speakers. Average F_0 decreases with increasing age until adulthood for both males and females, as shown in Fig. 1. The average drop for

BEST COPY AVAILABLE



(a)



(b)

Figure 1. Legend on following page.

Legend to Fig. 1 (previous page)

Figure 1. (a) The solid line shows mean fundamental frequency used by normally-hearing female speakers at different ages. Mean frequencies of 90 percent of speakers are expected to lie within the range indicated by dashed lines. (b) Same as (a), except for male speakers. The adolescent voice break can occur at different ages for different individuals. (Data from Fairbanks, 1940; Fairbanks, Herbert, & Hammond, 1949; Fairbanks, Wiley, & Lassman, 1949; Hollien & Paul, 1969).

females is roughly 75 Hz (from about 275-300 Hz to about 200-225 Hz) during the time from prepubescence to adulthood. For males the drop over the same period is likely to be about 150 Hz (from about 275-300 Hz to about 100-150 Hz), about 100 Hz of which may occur abruptly as a result of the adolescent voice break (Curry, 1940; Fairbanks, 1940). Several studies of voices of males between 20 and 30 years of age have placed mean or median F_0 for this group between 119 and 132 Hz (Hanley, 1951; Hollien & Shipp, 1972; Philhour, 1948; Pronovost, 1942). There is some evidence that, at least in the case of males, F_0 may again increase by as much as 30 or 40 Hz with advancing age (Hollien & Shipp, 1972; Mysak, 1959). Of course, for any given age, average F_0s span a considerable range, but about 90% of them would be expected to be within plus or minus 30-40 Hz of the population norms (Fairbanks, 1940; Fairbanks, Wiley & Lassman, 1949; Hollien & Paul, 1969). These ranges of average F_0 for female and male speakers are shown in Figs. 1a and 1b.

Pitch varies in the speech of an average speaker over a range of 1 to 1-1/2 octaves (Fairbanks, 1940). This variation is used to indicate stressed and unstressed vowels, to add emphasis to what is being said, and to carry information about the structure and meaning of a sentence. Stressed syllables are usually spoken with higher pitch than are unstressed syllables, although it may be more accurate to say that stressed syllables are accompanied by pitch change, either within the stressed vowel or in an adjacent vowel. The way in which pitch, amplitude, and duration interact to establish stress is still not fully understood.

Linguistic and semantic information is carried by pitch in several ways. A falling F_0 is used, for example, to signal the end of the final stressed vowel in a declarative sentence. At a

major syntactic break within a sentence, such a fall is followed by a rise in F_0 to indicate that the sentence is to continue. For certain types of questions, a rise in F_0 occurs in the final stressed syllable. Sentences which are ambiguous when printed (e.g., "Does he speak French or German?") can be spoken in an unambiguous way because, in part, of the intonation pattern that is imposed on the words. Also, messages beyond the words, sometimes subtle, sometimes poignant, can be conveyed by the way the utterance is inflected. Consider how many ways one can say "Isn't that nice?"

The difficulties that the deaf speaker has with pitch are of two general types: inappropriate average pitch and improper intonation. Intonation problems may in turn be divided into two major types: monotone voice and excessive or erratic pitch variation.

Several investigators have noted that deaf speakers are apt to have a relatively high average F_0, or to speak in falsetto voice (Angelocci, Kopp, & Holbrook, 1964; Boone, 1966; Engelberg, 1962; Martony, 1968). There is some evidence that this problem is greater for teenagers than for preadolescents, and particularly troublesome for adolescent boys (Boone, 1966). The results of the study by Angelocci, Kopp, and Holbrook (1964) suggest that not only are the fundamental frequencies of deaf speakers higher than those of hearing speakers, on the average, but the average F_0 for different speakers spans a wider range.

Deaf speakers often tend to vary the voice pitch much less than do hearing speakers, and the resulting speech has been described as flat or monotone (Calvert, 1962; Hood, 1966; Martony, 1968). A particular problem is that of inappropriate, or insufficient, pitch change at the end of a sentence (Sorenson, 1974).

A terminal pitch rise--such as that that tends to occur at the end of some questions--may be even more difficult for a deaf child to produce than a terminal fall (Phillips, Remillard, Bass & Pronovost, 1968). Deaf speakers who tend to produce each syllable with equal duration may also generate a similar pitch contour on each syllable. Such speakers may fail to indicate variations in stress either by changing the syllable durations or by modifying the pitch contours on the syllables. Thus, for example, a common error would be to fail to shorten an unstressed syllable and to lower the pitch on such a syllable.

That pitch problems vary considerably from speaker to speaker is well illustrated by the fact that, whereas insufficient pitch variation has been noted as a problem for some speakers, excessive variation has been reported for others (Martony, 1968). Such variations are not simply normal variations that have been somewhat exaggerated, but, rather, pitch breaks and erratic changes that do not serve the purpose of intonation.

It has been suggested that some of the unusual pitch variations that occur in the speech of the deaf may result from attempts by the speaker to increase the amount of proprioceptive feedback that he receives from the activity of producing speech. Martony (1968) and Willemain and Lee (1971) have observed that deaf speakers sometimes tend to begin a breath group with an abnormally high pitch, and then to lower the pitch to a more normal level. Willemain and Lee also noted that the average pitch of the deaf sometimes increases with the difficulty of the utterance. Inasmuch as the production of high pitch requires increased vocal effort (such as increased tension in the cricothyroid muscle and increased subglottal air pressure), they hypothesized that deaf speakers generate high-pitched tones as a way of providing

kinesthetic cues concerning the onset and progress of voicing. A similar conjecture was put forth by Angelocci, Kopp, and Holbrook (1964) who found that F_0 varied more from vowel to vowel when the vowels were produced by some deaf speakers than when produced by hearing speakers, while the reverse relationship held for first- and second-formant frequencies. These investigators attributed this type of abnormal pitch variation to efforts by the deaf speaker to differentiate vowels by varying F_0 and amplitude rather than the frequency and amplitude of the formants. "In physiological terms he is achieving vowel differentiation by excessive laryngeal variations with only minimal articulatory variations" (p. 169).

Some of the pitch variation from vowel to vowel may be the consequence of improper use of the muscles or use of inappropriate muscles or muscle groups in controlling the vowel articulations. These inappropriate muscle contractions may result in inadvertent tensing or slackening of the vocal cords, resulting in excessive variations in pitch.

Pitch has been described as a particularly difficult property of speech for deaf children to learn to control (Boothroyd, 1970). One possible reason for the difficulty is that deaf children may lack a conceptual appreciation of what pitch is (Anderson, 1960; Martony, 1968). Hearing people describe it in terms of a high-low dimension, but the description is somewhat arbitrary, and it is not clear that it is meaningful to an individual who has not had an opportunity to learn, by hearing, what "high" and "low" refer to in the auditory domain. A lack of intuitive grasp of the concept may help explain why deaf children often attempt to raise their pitch by increasing their vocal intensity (Phillips, Remillard, Bass, & Pronovost, 1968).

2.3 Velar Control

The velum or soft palate functions as a gate between the oral and nasal cavities. It is lowered to open the passage to the nasal pharynx when a sound such as one of the nasal consonants is made, which requires that the air be emitted through the nose. It is raised, thus sealing off the passage, for sounds that do not make use of nasal resonances, and in particular for those requiring the build-up of pressure in the mouth (obstruent consonants). Improper control of the velum has long been recognized as a source of difficulty in the speech of the deaf (Brehm, 1922; Hudgins, 1934). If the velum is raised when it should be lowered, the speech may be described as hyponasal; if it is lowered when it should be raised, hypernasality is the result. Miller (1968) has speculated that type of hearing loss may be a causative factor in connection of some nasalization problems. Hyponasality, he suggests, may be more prevalent among people with conductive loss than among those with sensorineural loss, because nasal sounds may appear excessively loud to the former, due to the transmittability of nasal resonances via bone conduction. Individuals with sensorineural loss, on the other hand, may welcome the additional cues provided by the nasal resonances, and therefore tend to nasalize sounds that should not be nasalized.

Nasality is often described as a "quality" problem because inappropriate velar control may give the overall speech a characteristic sound. In addition to affecting quality, however, inappropriate control of the velum can also lead to articulatory problems. There are three nasal consonants (m, n, and ng) which, in combination, account for about 11-12% of the occurrences of speech sounds in English (Denes, 1963; Dewey, 1923). A primary difference between the articulatory gestures involved in the production of these nasal

consonants and those that are used to produce the three stop consonants (b, d, g) is that the velum is lowered in the former case and raised in the latter. If the velum is raised when it should be lowered, or lowered when it should be raised, confusions between pairs of these sounds may occur. Not surprisingly, such substitutions are found in the speech of the deaf (Stevens, Nickerson, Boothroyd, & Rollins, 1974).

Learning appropriate velar control may be particularly difficult for a deaf child for two reasons: (1) raising and lowering the velum is not a visible gesture and therefore not detectable by lipreading; (2) the activity of the velum produces very little proprioceptive feedback. Normally-hearing persons are relatively insensitive to the activity of this part of the articulatory apparatus, and may be quite unable, without practice, to manipulate it as an act of conscious control, say while making a steady vowel sound. Obviously, the movement of the velum must be timed fairly accurately when producing words with abutting nasal and stop consonants if the appropriate sounds are to be produced and the resulting speech is to be fluent. Deaf speakers often have considerable difficulty producing such clusters (Stevens, et al., 1974).

Improper velar control is difficult to judge subjectively, in part because the distinctive perceptual features of nasalization are not clearly defined and in part because the perception of nasality may be affected by factors in addition to the activity of the velum. A deliberate constriction of the nasal pathways, for example, can modify the resonant characteristics of nasal consonants and adjacent vowels, thus producing a type of "nasal speech" which does not necessarily involve improper velar control. Also, some researchers have suggested that the perception of nasality may be influenced by such factors as malarticulation, pitch variations,

and speech tempo (Colton & Cooker, 1968). For these reasons, objective measures that correlate with velar activity are of considerable interest to investigators of speech. Acoustic properties of nasal sounds that have been investigated include shifted and "split" first formant (Fujimura, 1960; House, 1961) and enhanced amplitude of the lowest harmonics (Delattre, 1955). Attempts to detect nasalization directly have included the measurement of the flow of air through the nose (Lubker & Moll, 1965; Quigley, Schiere, Webster, & Cobb, 1964), the acoustic energy radiated from the nostrils (Fletcher, 1970; Shelton, Knox, Arndt, & Elbert, 1967), and the vibration on the surface of the nose (Holbrook & Crawford, 1970: Stevens, Kalikow, & Willemain, 1974).

Procedures have not yet been developed for quantitatively assessing the severity of nasalization problems. Indeed, normative data are lacking that would provide the necessary baseline measures in terms of which deviance could be judged. Moreover, what may be more important--at least for intelligibility--than the overall nasality of an individual's speech is the difference in the degree of nasalization of sounds that should be nasalized and those that should not be, and the adequacy of the velar adjustments that are required in order to produce nasal consonants in the context of other sounds. Stevens, Nickerson, Boothroyd, and Rollins (1974) have defined for this purpose an index based on detection of vibration by an accelerometer attached to the surface of the speaker's nose. The index is intended to indicate how well the speaker differentiates nasal consonants and nonnasal vowels in running speech. It is defined as the difference between the average amplitude of the accelerometer signal (in decibels) for nasal consonants and the amplitude for vowels that should be produced without nasalization. Measurements obtained from normally-hearing speakers produced values of this index in the range

10-20 dB. Values close to zero would suggest a failure to differentiate nasal from nonnasal sounds. Such failure could result from either excessive hyper- or hyponasality.

2.4 Articulation

Articulatory problems of a variety of types have been identified. Failure to develop certain phonemes, failure to differentiate between others, substitution of one sound for another, use of the neutral schwa /ə/ (as in about) as a general-purpose vowel, and other distortions of pronunciation of various sorts are all articulatory difficulties that are encountered in the speech of the deaf. Unfortunately, the information that exists concerning these problems is fragmentary and not easily integrated into a self-consistent whole. No large-sample study has been conducted for the express purpose of cataloging the various articulatory problems that are found in the speech of the deaf, or of determining either the prevalence of individual problems or their relative importance vis-a-vis intelligibility and speech quality. Several investigators have reported specific articulatory difficulties that they have observed among particular groups of deaf children, and it is these reports that will be summarized here. It will be apparent that an intensive study of a large enough sample of the speech of the deaf to provide some reliable prevalence data, and a concerted attempt to relate articulatory problems to intelligibility and quality measures, could increase greatly our understanding of the speech problems of the deaf.

The failure to produce appropriate vowel sounds has been noted as a problem by several investigators (Angelocci, Kopp, & Holbrook, 1964; Boone, 1966; Hudgins & Numbers, 1942). The problem may take the form of a failure to differentiate one vowel sound from another, or that of producing diphthongs in place of vowels. Typically, vowel errors tend to involve spectrally similar sounds (Smith, 1973).

Inasmuch as the formant frequencies--and perhaps especially F_2 (Boothroyd, 1972; Licklider & Pollack, 1948; Thomas, 1968)--apparently provide the information that is needed to distinguish among different voiced speech sounds, one might guess that the speech of the deaf would tend to show some deficiencies in this respect. Some such deficiencies have been noted. Boone (1966), for example, found that the second-formant frequency tended to be lower for deaf than for hearing children, a fact which he attributed to the tongue being too far back toward the pharyngeal wall. (The inappropriate tongue position was also considered by Boone to be responsible for the "cul de sac" resonance that has been ascribed to some deaf speakers.) This observation is consistent with Mangan's (1961) identification of faulty front vowel production as one of the major contributors to the errors that listeners made in transcribing a list of 50 PB words read by deaf speakers.

Angelocci, Kopp and Holbrock (1964) have also focused on formant frequencies in a comparison of the speech of another sample of deaf and hearing children. In this case, the range of mean values of F_1 was much smaller for the deaf than for the hearing children, and the difference between F_0 and F_1 was smaller for the former group. The range of the means of F_2 was also smaller for the deaf children than for those with normal hearing, and the dependence of the frequency and amplitude of F_3 on which vowel was being spoken seemed to differ considerably for the two groups. As we have already noted in the discussion of pitch and intonation, F_0, in contrast to F_1 and F_2, varied more with vowels for the deaf than for the hearing speakers.

Angelocci, et al. present several coordinate plots of F_2 against F_1, showing the scatter of these coordinates for each vowel for both deaf and hearing speakers. What is clear from the plots

is that the degree of overlap among the areas representing different vowels is much greater in the case of the speech samples obtained from the deaf. For the hearing speakers, the easiest and most difficult vowels to identify were, respectively, /i/ (98%) and /æ/ (48%); /æ/ was frequently misidentified as /ɛ/. The best and worst vowels produced by the deaf speakers were, respectively, /u/ (46%) and /ə̂/ (21%).

Several investigators have claimed that as a rule deaf speakers are better at producing consonant sounds than at producing vowel sounds (Huntington, Harris, Shankweiler, & Sholes, 1968; Joiner, 1922; Jones, 1967; for a counter example, see Nober, 1967). Nevertheless, many difficulties associated with consonant sounds have been noted. Stewart (1969) identifies the production of fricatives and affricates as one difficulty and points out that /s/ and its voiced cognate /z/ are often omitted altogether, particularly from the final syllable position. Other investigators have also identified /s/ as a special difficulty (Borrild, 1968; Brehm, 1922; Nober, 1967), as well as the failure to distinguish between /s/ and /š/ (shoe). This is perhaps not surprising inasmuch as most of the energy of the /s/ sound is concentrated at the high end of the frequency range, where the hearing deficit of many hearing-impaired individuals, particularly those with sensorineural loss, tends to be most severe. (Apparently, even people with close-to-normal low-frequency hearing may have difficulty producing good sibilants if they have severe hearing loss above about 1 kHz [Miller, 1968] or speech difficulties not related to hearing [Irwin, 1966].) Also, the articulatory gestures involved in producing /s/ and /š/ are relatively invisible.

Failure to distinguish between voiced and voiceless consonants is another problem that has been noted (Calvert, 1961, 1962; Mangan, 1961; Smith, 1973). One form of this problem is the "surd-sonant

error" (Calvert, 1962, 1964; Mangan, 1961) in which intended voiced plosives are perceived as voiceless plosives, or the reverse. Calvert (1962) measured the durations of closure and release periods of consonants and found that when a plosive was intended to be unvoiced (e.g., p, t) and was heard as voiced (e.g., b, d), the duration of the release period was about the same as that of the voiced consonant when produced by a hearing speaker. Similarly, when a voiced consonant was intended and its unvoiced cognate was perceived, the measured duration of the release period was appropriate to the perceived form. Another form of the voice-voiceless problem is continuous phonation, a defect which, according to Millin (1971), is observed in a "sizeable segment" of the hearing-impaired population and contributes significantly to reducing the intelligibility of the speech. Millin notes that the fact that the phonation is continuous is not necessarily perceived immediately; listeners are more likely to be aware of severe misarticulation of phonemes as a result of it, but not of the underlying continuous phonation itself.

Another articulatory difficulty that has been observed is the omission of arresting and releasing consonants (Hudgins & Numbers, 1942). In particular, Stewart (1969) has noted as problems both the introduction of intrusive stop elements into the pronunciation of fricatives and the omission of stop elements when they should be there. The pronunciation of sheep as cheap is an example of the first type of problem, and the pronunciation of chair as share is an example of the second. He notes that intrusive stop elements can give the speech a somewhat "clipped" quality if they occur frequently enough.

There is some evidence from electromyographic data that the articulatory behavior of deaf speakers is more nearly like that of

hearing speakers with respect to lip movements than with respect to tongue movements (Huntington, Harris, Shankweiler, & Sholes, 1968), and, consequently, labial consonants produced by the deaf tend to be more intelligible than lingual consonants and vowels. This could be due either to the greater visibility of lip movements, or to the possibly greater inherent complexity of tongue gestures; however, Huntington et al. concluded agairst attributing the difference to the greater difficulty of tongue movements on the grounds that a similar greater intelligibility of labial sounds was not found for hearing individuals who had speech difficulties stemming from central nervous-system disorders. The importance of the relative visibility of articulatory gestures in determining the ease with which the deaf learn to produce specific sounds is further suggested by Guttman, Levitt, and Bellefleur's (1970) finding of a positive correlation between the quality of a speaker's articulation and his lipreading ability. In the same vein, Levitt (1974) has pointed out that the speech production errors that were documented by Smith (1973) show patterns of confusions among phonemes that are similar to those found by Erber (1974) in studies of lipreading.

Perhaps the most compelling evidence concerning the importance of visibility as a determinant of articulatory competence has been reported by Nober (1967). He found that when consonants, classified in terms of place of articulation, were rank-ordered in accordance with the relative frequency with which they were correctly articulated by the 46 deaf children in his study, the resulting order (from best to worst: bilabial, labiodental, glottal, linguadental, lingua-alveolar, linguapalatal, linguavelar) was very similar to the order that would represent relative visibility. (Nober also reported the following order for articulatory competence in terms of manner of articulation, again from best to worst: glides, stops, nasals and fricatives.) That visibility is not the only factor,

however, is suggested by the fact that several studies of consonant articulation problems arising from causes other than hearing loss (retardation, cleft palate) have shown similar trends with respect to the rank ordering of consonants in terms of how problematic they appear to be. Such consistency, Nober notes, is suggestive of the importance of maturational factors, and of differences in the inherent difficulty of articulating and of discriminating different phonemes.

It is important in discussing articulatory problems to distinguish between the ability to produce appropriate individual speech sounds in isolation and the ability to combine those sounds in such a way as to produce fluent speech. Deaf children often have the former ability, but not the latter (Borrild, 1968; Jones, 1967). Difficulties in executing smooth transitions between speech sounds (e.g., consonant-vowel transitions) (Jones 1967) and malarticulation of compound and abutting consonants (Hudgins & Numbers, 1942) are examples of articulation problems that probably affect fluency detrimentally. These difficulties may also help to account for the finding that consonants occurring in initial sound position tend to be articulated better than those occurring in medial position, which in turn tend to be better than those in final position (Nober, 1967). Durational aberrations in transitional sounds (Jones, 1967) would be especially detrimental to smoothly flowing speech, as would many of the timing deficiencies discussed in the section of this paper dealing with timing problems.

A few investigators have attempted to determine how the spontaneous development of speech sounds differs between deaf and hearing children. Some have suggested that during the first year of life deaf and hearing children do not differ greatly in their spontaneous vocalizing and babbling (Carr, 1953). However, recent

findings have provided evidence of differences in babbling of deaf and normal-hearing infants as early as 22 weeks (Murai, 1961; Malvilya, 1970 [cited in Menyuk, 1972]). In general, the relatively "easier" sounds appear to be the more prevalent, e.g., middle as opposed to extreme front and back vowels, and voiced labial consonants as opposed to unvoiced and lingual consonants (Heider, Heider, & Sykes, 1941; Neas, 1953 [cited in Carr, 1964]), although there is some evidence that most of the sounds that occur in standard English can be found in the spontaneous vocalizations of deaf children (Sykes, 1940; Carr, 1953; Fort, 1955 [cited in Carr, 1964]). According to Carr (1953), children's speech shows a greater frequency of vowel sounds than of consonant sounds (see also Neas, 1953), of front vowels than of back vowels; and, at least in the case of deaf children, front consonants occur more frequently than back consonants. In addition, voiced consonants are more prevalent than their unvoiced cognates. Carr points out that the differences between the speech sounds of deaf and hearing children become increasingly more apparent with increasing age, and suggests that the differences are best characterized by saying that the spontaneous development of sounds by deaf children does not continue much beyond one year. He suggests that the higher frequency of front consonants among deaf children may be attributed to the greater visibility, and, hence, imitatability of the articulatory gestures involved. It is believed that the articulatory maturation of children with normal hearing is complete by about the eighth year (Templin, 1957).

2.5 Voice Quality

It seems to be generally agreed that deaf speakers have a distinctive voice quality (Bodycomb, 1946; Boone, 1966; Calvert, 1962); however, what exactly is meant by voice quality is not entirely clear. More specifically, the question arises whether

it is more appropriate to think of the speech of the deaf as having a distinctive voice quality, or as having a variety of qualitative properties which characterize the speech of different individuals to different degrees. Unfortunately, data relating perceived quality to the acoustic properties of speech are sparse, almost to the point of nonexistence.

The term "quality" is sometimes used in contrast with "intelligibility," the idea being that the quality and intelligibility of speech may vary somewhat independently. Sometimes the term also appears to be used to connote steady-state, as opposed to dynamic, properties of speech. In this sense, hypernasality and hyponasality might be considered qualitative properties of speech, whereas rhythm and timing aspects probably would not. Another example of a steady-state characteristic that could contribute to qualitative distinctiveness is the "cul de sac" resonance described by Boone. Still another is breathiness (Hudgins, 1937; Peterson, 1946; Scuri, 1935, reviewed by Hudgins, 1936), a characteristic that Hudgins attributed, in large measure, to inappropriate positioning of the vocal cords and poor control of breathing during speech. In particular, too large a glottal opening may be produced by failure to close properly the vocal folds: "the result is a large expenditure of air and a voice of poor quality" (Hudgins, 1937, p. 345).

This list of problems relating to speech quality could be extended considerably. Calvert (1962), in fact, was able to find 52 different adjectives that had been used as descriptors of deaf speech in the literature. When fifteen teachers of the deaf were asked to select from these 52 words those that they considered to be the most accurate, the words most often chosen were "tense," "flat," "breathy," "harsh," and "throaty."

Calvert (1962) also attempted to determine empirically whether in fact the speech of the deaf is distinguishable on the basis of quality from that of speakers with normal hearing. He had teachers of the deaf attempt to determine by listening whether recorded speech sounds (vowels and diphthongs in isolation, nonsense syllables, words, and sentences) had been produced by profoundly deaf speakers, speakers imitating deaf speech, speakers simulating harsh and breathy voice, or normally-hearing speakers. Isolated vowels, from which onset and termination characteristics had been clipped, could not be distinguished as to source; however, the sources of the sentences were identified with 70% accuracy. Calvert concluded that deaf voice quality is not identified on the basis of relative intensity of the fundamental frequency and the harmonics alone, but on the dynamic factors of speech such as the transition gestures that change one articulatory position into another.

Although it is questionable whether inappropriate "loudness" or "volume" of speech is best thought of as a quality deviation, it is perhaps as reasonable to mention it here as under any of the other categories in terms of which this review is organized. The problem, which has been noted by several investigators (Carhart, 1970: Martony, 1968; Miller, 1968), may take several forms: voicing may be too soft, or too loud, or the volume may vary erratically. Miller (1968) points out that the way in which the volume of a speaker's voice is affected by hearing loss may depend on the nature of the impairment. An individual with a sensorineural loss may tend to speak in an abnormally loud voice because he does not receive feedback via bone conduction, whereas an individual with a conductive loss may tend to speak very softly, because his own voice, which he may hear via bone conduction, may appear very loud as compared with the speech of persons with whom

he is talking. Carhart (1970) advocates that deaf people be trained to talk at each of four or five general levels of loudness, and to shift from one to the other, depending on kinesthetic cues and reactions from listeners to judge the appropriateness of the level at which they are talking at any given time.

How important voice quality is for intelligibility is really not known. One can find a variety of views on this issue in the literature. Peterson (1946), for example, considers voice quality to be relatively unimportant as a determinant of intelligibility. Adams (1914), on the other hand, points out that while it may have little effect on intelligibility in a technical sense, it can play a very important role in determining whether what a deaf speaker is saying will in fact be understood by an unfamiliar listener. She claims that people who are unfamiliar with the deaf may find their speech so disagreeable when they first encounter it that, even if it is quite adequate for effective communication, the listener may not make the effort necessary to understand it.

2.6 Interrelatedness of Problems

As has already been noted, the topical organization of the foregoing discussion of problems encountered in the speech of the deaf was used as a writing convenience, and, while it is not an unreasonable organization, there are doubtlessly others that would serve as well. It is important to emphasize, however, that any problem taxonomy has an element of arbitrariness about it. And the impression that any taxonomy can create that the subject is really partionable into several independent classes is, in this case, certainly false. The problems that have been discussed under separate headings interrelate in many ways.

The importance of accurate timing at the phonetic level for correct articulation is intuitively apparent, for example, as is the necessity for accurate velar control.

Calvert (1962) makes the point that the types of durational distortions that impair the intelligibility of the speech of the deaf (e.g., extension of unstressed vowels, fricatives and closure periods of plosive consonants) may also contribute to perceived speech quality. He notes that deaf speakers are more easily distinguished from speakers with normal hearing, the greater the articulatory complexity of the utterance, and concludes that distortions in phoneme durations may be significant determinants of what is commonly called "deaf voice."

Hudgins (1934, 1936, 1937, 1946) has extensively documented the interrelationship between the problem of inappropriate control of breathing during speech and that of poor timing and rhythm.

Peterson (1946) speaks of the lack of pitch variation as one of the three major quality problems in the speech of the deaf. (The other two that he identifies are breathiness and nasality.) We have classified lack of pitch variation, or monotone speech, as a pitch control problem, but it clearly could easily be discussed under the topic of speech quality. Furthermore, the inappropriate laryngeal posture that seems to be a concomitant of breathy voice quality undoubtedly has an influence on the control of laryngeal muscles that produce changes in pitch.

Colton and Cooker (1968) have suggested that the perception of nasality may be influenced by such factors as articulatory errors, pitch variations and slower than normal tempo.

While conceptually distinct, the problems of volume control and pitch control are probably closely related in practice. There is some indication that a deaf child has difficulty gaining separate control over volume and pitch: often he tends to increase vocal effort when trying to increase pitch (Phillips, Remillard, Bass, & Pronovost, 1968).

That pitch, volume and timing are intimately interdependent as determiners of stress patterns is well known, but exactly how they relate is not. In some cases a change of emphasis might be indicated by a change in any one of these properties. In other cases, however, changes probably have to occur together. It is evident, for example, that a vowel that must carry a falling pitch contour must be lengthened to accommodate this contour.

The list of factors that interrelate could be lengthened; indeed, one might argue persuasively that each of the problems that has been discussed here is related in some way to each of the others. But perhaps the point has been made. While an analytical approach is undoubtedly necessary in order to make the task of studying deficient speech tractable, a problem-by-problem approach to training may be bound to yield only limited success. Perhaps the most distinguishing characteristic of speech is its integrity, and it may be that training techniques will continue to produce disappointing results until methods are developed that will provide a child with an ever-present visual or tactual representation of his own and other people's speech that is as rich as the representation that the hearing person gets by ear. But exactly what information should be represented in such a display and how that information should be encoded, are questions that have yet to be answered. And until they are answered, an analytic approach to training is probably the only path that is open.

3. TRAINING PHILOSOPHY

We have found it helpful to distinguish four major aspects of the problem of teaching speech to the deaf:

1. diagnosis of speech deficiencies,
2. establishment of training objectives,
3. specification of training procedures, and
4. evaluation of progress.

While each of these aspects is essential to a well-rounded speech training program, it will be apparent from what follows that we consider objective diagnosis to be basic. Given an adequate diagnostic procedure, much of the rest follows: without such a procedure it is questionable whether the other desiderata can be realized at all.

3.1 Diagnosis

In diagnosing a child's speech problems it is not enough to say that the speech is unintelligible, or that it does not sound natural. The goal should be to specify as precisely and quantitatively as possible the nature of the speech deficiencies and how these deficiencies contribute to the lack of intelligibility or poor quality of the speech. What is needed is a diagnostic profile that is to speech what the audiogram is to hearing.

Such a profile has yet to be developed. Moreover, the development of a diagnostic procedure that produces a demonstrably valid and reliable assessment of one's level of competence with respect to the various skills that make for speech proficiency must undoubtedly await a fuller understanding of both "normal" and defective speech. However, until such a procedure is developed, the other aspects of speech teaching will necessarily be based on less than

solid ground. Therefore, we make the assumption that even a crude diagnostic procedure is better than none, provided that (a) its limitations are recognized and (b) that it is viewed as a point of departure for the development of more precise and useful techniques.

It seems likely that any diagnostic profile that is to be reasonably comprehensive must contain information concerning each of the problem categories that have been discussed in Section 2. The findings reviewed in that section suggest the following measurements as candidates for a diagnostic profile.

Timing and rhythm

- Speech rate (rate of word, syllable or phone emission)
- Ratio of durations of stressed and nonstressed syllables (probably distinguishing stressed syllables in final and non-final word and sentence positions)
- Ratio of pause time to speaking time in utterances of varying complexity and rhythmic structure
- Relative frequency of pauses in different contexts (between sentences, between phrases, within phrases)
- Durations of specific speech sounds (e.g., fricatives, closure periods of stop consonants)
- Timing of coarticulation and transition events (e.g., timing of voice onset relative to release of the closure period of plosive consonants, onset of nasalization for nasal consonants, timing of glide from one vowel sound to another in production of diphthongs, transitions between fricative or nasal consonants and vowels)
- Rhythm (subjectively judged)

Pitch and intonation

- Average F_0
- Magnitude of F_0 fall (in octaves) after F_0 maximum in final stressed vowel in selected declarative sentence(s)

- Magnitude of F_0 rise on sentence-final stressed vowel in selected interrogative sentence(s)

- Magnitude of F_0 rise from initial unstressed vowel to initial stressed vowel in selected sentence(s)

- Variability of F_0 for selected utterance(s)

- Acceptability of intonation pattern on selected sentences (subjectively judged)

Nasalization

- Nasalization level (measured from an accelerometer attached to the nose) while sustaining /m/ sound

- Nasalization level while sustaining each of several vowel sounds

- Peak nasalization levels when saying selected monosyllables such as ma, pa, no, toe, song, sock

- Nasalization index for selected sentences and phrases (see Section 2.3, and Stevens, Nickerson, Boothroyd, & Rollins [1974] for a definition of this measure)

Articulation

- Spectral differentiation among vowels, especially /a/, /i/, and /u/

- Spectral distinctiveness of fricatives, especially /s/ and /š/

- Differentiation between voiced consonants and their voiced cognates

- Differentiation between nasal consonants and nonnasals whose production depends upon the same place of articulation

- Appropriate production of consonant clusters and blends

- Smoothness of glide from one vowel sound to another in the production of diphthongs

Quality

- Measure of breathiness

- Measures listed under other categories insofar as they relate to speech quality

Missing from this picture are normative data in terms of which the normalcy or deviance of measures such as these can be determined. In some cases, norms can be inferred from the literature; in other cases, they will have to be developed empirically. If a measurement profile can be developed that can be demonstrated to be descriptive of those aspects of speech that are critical to intelligibility and quality, it will undoubtedly be desirable to develop norms for the material in the diagnostic procedure itself. This would mean applying the procedure to a large and representative sample of speakers in order to obtain accurate estimates of the way in which the various measures are distributed in the general population. The magnitude of such an undertaking is sufficiently great, however, to be justified only after considerable progress has been made in establishing what the measures should be.

3.2 Training Objectives

The ultimate training objective, of course, is intelligible, fluent speech. Such an objective is not very helpful to the speech teacher, however. What is needed are some interim objectives that satisfy certain criteria, among which are the following:

Assessability. Interim objectives must be defined in such a way that one may determine whether or not they have been attained at any particular time. To the extent possible, they should be quantitative. Ideally, they should be expressed in terms of the properties and measurements that comprise the diagnostic profile. Examples of assessable objectives relating to the diagnostic measurements listed above would include: modifying average F_0 to bring it within a specified range, increasing the ratio of durations of stressed and unstressed syllables, increasing syllable emission rate, increasing the amount of pitch fall on sentence-final syllables, improving quantitative measures that relate to adequacy of velar control.

Feasibility. Interim objectives should be realistic in the sense that there is a reasonable expectation of achieving them. This implies the need for graded objectives that are applicable at different stages of speech proficiency. It also reinforces the notion that an effective diagnostic procedure is the sine qua non of a comprehensive training program. What is a feasible training objective for one individual may be unrealizable for another. And the determination of what is feasible and what is not must be made in the light of an accurate assessment of an individual's current level of speech proficiency and his specific difficulties.

Validity. Interim objectives should be such that achievement of them does in fact facilitate realization of the long-term goal of improving speech intelligibility and quality. It is important to recognize, however, that achievement of any particular interim objective will not necessarily, in and of itself, improve intelligibility or quality. It could diminish either or both. Improvement of overall rhythm could, for example, have the immediate effect of degrading articulation and thereby decreasing intelligibility. However, working toward an interim goal in such a case might still be justified providing that it is clear that the long-range gains more than offset the short-term impairment. How one makes that determination in any particular case, given the current level of understanding of speech production and recognition, is not apparent. However, the existence of a widely used diagnostic profile could itself help considerably to develop the data necessary to solve the problem. If one had a large set of speech samples, each of which had been assessed with respect to intelligibility and overall quality, and for each of which a diagnostic profile had been obtained, one could begin to determine the relative importance of various deficiencies as determinants of intelligibility and quality.

These comments with regard to training objectives must apply not only to specific vocal exercises and utterances that are rehearsed and produced in a training situation. Appropriate objectives must be formulated to encourage carryover of the skills learned in the tutorial sessions to spontaneous speech. In fact, it frequently happens that certain specific speech skills can be acquired rather rapidly in the tutorial sessions, but the carryover of these skills to everyday speech is a long and difficult process.

3.3 Training Procedures

Training procedures should be designed to reduce the distance between where one is and where one would like to be. It seems fairly apparent that such procedures cannot be developed, or at least not tailored to the needs of individual students, with any assurance of efficacy in the absence of adequate diagnostic data and concrete training objectives. This is not to suggest that the specification of individualized training procedures should be a trivial task, given acceptable solutions to the diagnosis and training-objective problems, but rather to suggest that it is probably an impossible task if these other problems are not at least partially solved.

In Section 4 of this report we present numerous examples of how the displays that the system can generate can be used in training. In particular, many suggestions are made concerning exercises that may be used in an effort to realize specific objectives. It is important, of course, that these suggestions be applied in ways that are consistent with an individual child's diagnosis and the particular training objectives that have been established for him.

3.4 Evaluation of Progress

The fourth problem, like the third, is greatly diminished by the existence of acceptable solutions to the first two. To the extent that training objectives are expressable in quantitative form, progress vis-a-vis these objectives can be assessed in terms of specific quantitative measurements. Progress in a more global sense can be evaluated in terms of changes in the overall speech diagnostic profile. It would, of course, be necessary to establish norms with respect to the various measurements comprising the profile and such norms would have to reflect different speech characteristics that are found in different groups of speakers. For example, one would want to express pitch norms in terms of age and sex. Given the establishment of such norms, an ultimate goal of training vis-a-vis objectively measurable aspects of speech would be to produce speech for which the values of the parameters represented on the profile would fall within the appropriate normative range with respect to all measurements.

In this section we have been discussing philosophy, and attempting to articulate a point of view concerning speech training. This point of view recognizes four major aspects of the speech-training problem: diagnosis, training objectives, training procedures, and performance evaluation. We recognize the danger of oversimplification in any conceptual scheme of this sort. And we certainly do not wish to suggest by this particular conceptualization that we believe speech training can really be reduced to a by-the-numbers approach. We do, however, want to insist that each of the four factors mentioned is, or should be, an important aspect of speech training, to argue that an effort to develop a program that incorporates each of those aspects explicitly and in an integrated way should be made, and, finally, to suggest that diagnosis is in some sense the fundamental problem.

4. USE OF DISPLAYS FOR SPEECH DIAGNOSIS AND TRAINING

The purpose of this section is to discuss and illustrate ways in which some of the displays that have been implemented on the BBN system can be used to facilitate diagnosis of an individual child's speech problems and to help provide training aimed at remediation. The discussion here is limited to displays that are implemented on the system as it currently exists. There is always the possibility, of course, of adding to the system's capabilities, and, in particular, of programming new displays for either diagnostic or training purposes as needs for such additions are identified.

The section is organized in terms of the problem areas that were discussed in Section 2. It will be clear, however, that it is often difficult to maintain the separability of the different problem areas, both because of the fact that many of the displays provide information on several aspects of an utterance and also because of the inherent interrelatedness of the problems on which we have focused. Thus, while a separate subsection is devoted to each of the major topics, we do not hesitate to violate this topical arrangement somewhat when it appears reasonable to do so in order to make optimal use of the illustrative material that is presented.

The system itself and the displays that it is capable of generating have been described elsewhere (Nickerson & Stevens, 1973; Nickerson, Kalikow, & Stevens, 1974). Details concerning the mechanics of loading display programs, setting display parameters, and so forth, may be found in Rollins, Kalikow, and Nickerson (1974).

Finally, we wish to stress that the suggestions in this section are suggestions only. They are offered as working hypotheses about how best to proceed, and as points of departure for future research.

4.1 Timing and Rhythm

The diagnosis and training of speech timing and rhythm require the assessment of the temporal properties of several types of utterances, ranging from monosyllables through short phrases and sentences that contain syllables with various patterns of stress.

The first step in the diagnosis of timing problems is to measure the length of a single-syllable consonant-vowel utterance produced by the student. The student should be able to limit the length of the voicing in this syllable to 500 msec. or less, with a reasonably abrupt rise and fall in loudness. In assessing the timing for monosyllables and for other more complex utterances, it is appropriate to select words containing consonants and vowels that do not present serious articulation problems, i.e., articulation problems that might interfere appreciably with timing. Examples of time-plot displays for the syllable paw produced with acceptable duration and loudness contour are shown in Fig. 2. The solid horizontal line indicates the presence of voicing, and the contour is a measure of intensity of the signal (with a particular kind of frequency weighting). This display has been called the voicing-loudness (VL) display. The kinds of problems that are likely to be encountered at this stage include: (1) too long a duration of voicing (Fig. 3), (2) a gradual reduction in amplitude toward the end of the vowel indicating a possible breathiness in the vowel offset (Fig. 3a), and (3) too abrupt an onset of loudness of the vowel, as indicated by a brief peak in loudness at the beginning of the vowel (Fig. 3c).

The gradual reduction in amplitude at the end of the syllable, possibly accompanied by a cessation of the voicing line well before the final drop in loudness (as in Fig. 3a) is often observed when the utterance is judged to have a breathy termination. Under these

BEST COPY AVAILABLE

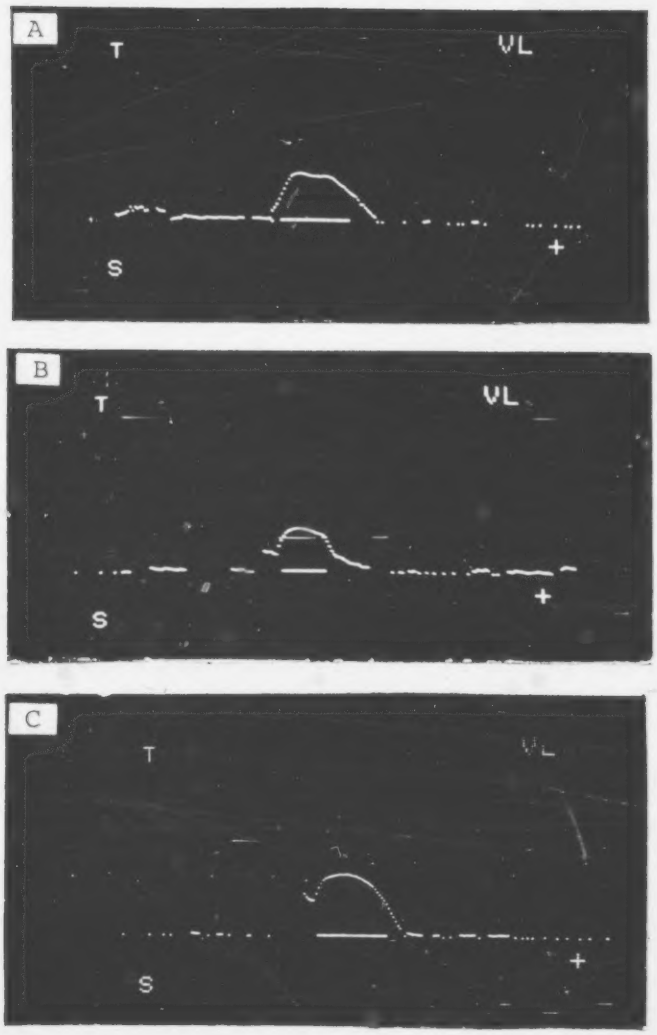


Figure 2. Voicing-loudness (VL) display for the monosyllabic word paw produced in an acceptable fashion by three different normally-hearing speakers. The total width of the display is about 2 sec. in this and in all subsequent time-plot displays.

BEST COPY AVAILABLE

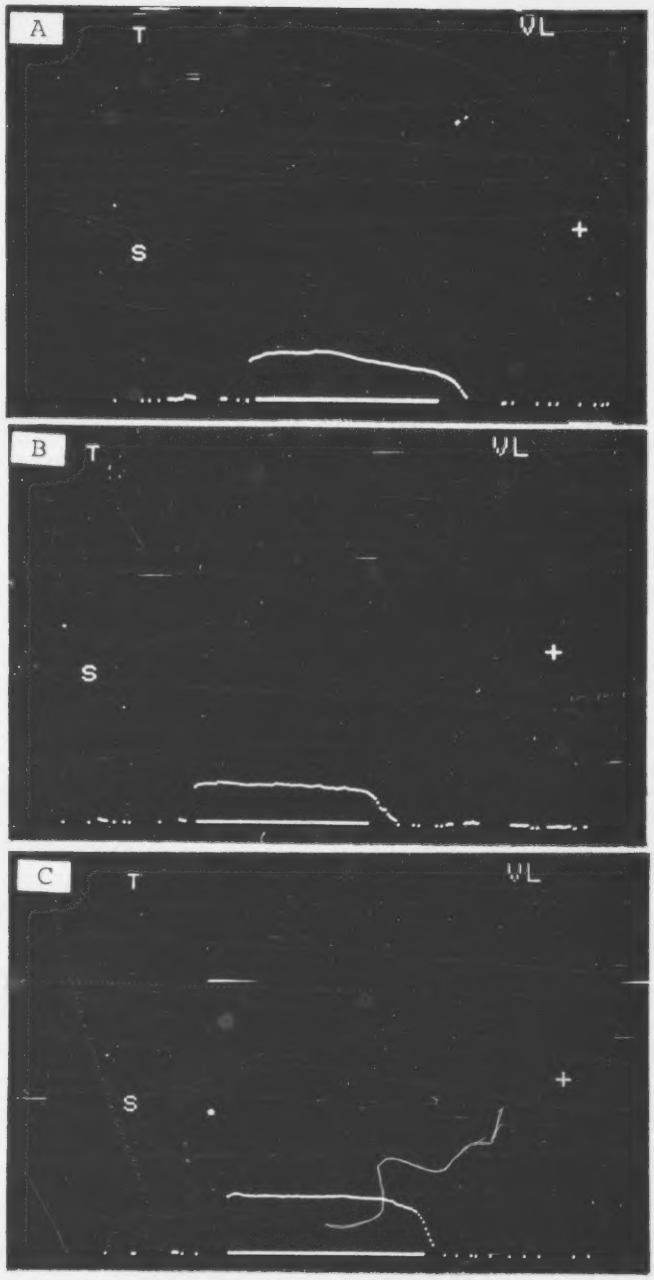


Figure 3. VL display for the monosyllabic word paw produced in an unacceptable way by three different deaf students. See text.

circumstances, the gradual decay is probably a consequence of a gradual and premature abduction of the vocal cords toward the end of the syllable--a premature return of the vocal-cord configuration to a position appropriate for normal breathing. The abduction maneuver is accompanied by a decreased amplitude of vocal-cord vibration, an increase in airflow, and an increase in the amplitude of turbulence noise generated at the glottis, resulting in the impression of breathy voice. Normally-hearing speakers often produce this kind of aspiration in the final syllable of an utterance, particularly if the syllable is an open syllable (i.e., no final consonant), but this gesture is much more marked for many deaf speakers, and leads to an unnatural sound.

The overly abrupt syllable onset (as in Fig. 3c) is probably a consequence of too great an effort in producing the initial consonant in the syllable. This increased effort--possibly reflected in a raised lung pressure--leads to a strong consonantal release, and a peak in loudness at the vowel onset. The rather "forced" or "tense" sound that results may in some cases stem from a preoccupation with the training of consonant articulation in isolated monosyllables.

For a student who produces a monosyllable with a time pattern that is too long or that has an inappropriate decay in loudness, the emphasis in training should be on forming a proper termination of the syllable. In the case of an open syllable, the voicing should be terminated relatively abruptly, avoiding the slow decay in amplitude noted above.

After monosyllables are assessed, the next step in the diagnosis is to examine the timing of simple two-syllable utterances with a

reduction in stress for either the first or the second vowel. Timing can be measured more easily with these two-syllable utterances if the consonant (or consonants) that separates the two syllabic nuclei is voiceless, since a break in voicing can then be observed, and the relative durations of the voiced portions of the two syllables can be determined. If there is no break in voicing, as in an utterance like the man, or sinner, then the relative durations of the syllables must be evaluated from the positions and durations of the peaks in the loudness contour.

For an utterance with an initial unstressed syllable (such as a car, the tree), the duration of voicing for the second (stressed) syllable should be 4 to 8 times the duration for the unstressed syllable (Nickerson, Stevens, Boothroyd, & Rollins, 1974). Examples of normal variation are given in Fig. 4. Usually the amplitude of the unstressed syllable (as indicated by the loudness contour) should be less than that of the stressed syllable. (There may be occasions when this amplitude relation does not apply, however, particularly when the stressed syllable contains a high vowel such as /i/ or /u/, which has an inherently lower amplitude than a non-high vowel such as /a/, /e/, /o/.) The VL display for the stressed vowel in these two-syllable utterances should satisfy the requirements indicated above for a monosyllable. It is important to note, however, that when the same stressed syllable occurs within an utterance and is not the final syllable before a pause, its duration should be less than the duration in utterance-final position.

For a two-syllable utterance with stress on the initial syllable (such as puppy, sister), the temporal pattern is quite different. Here the duration of voicing in the initial stressed syllable should be 1/2 to 3 times the duration of voicing in the unstressed syllable. Furthermore, both the stressed and the unstressed syllables should be

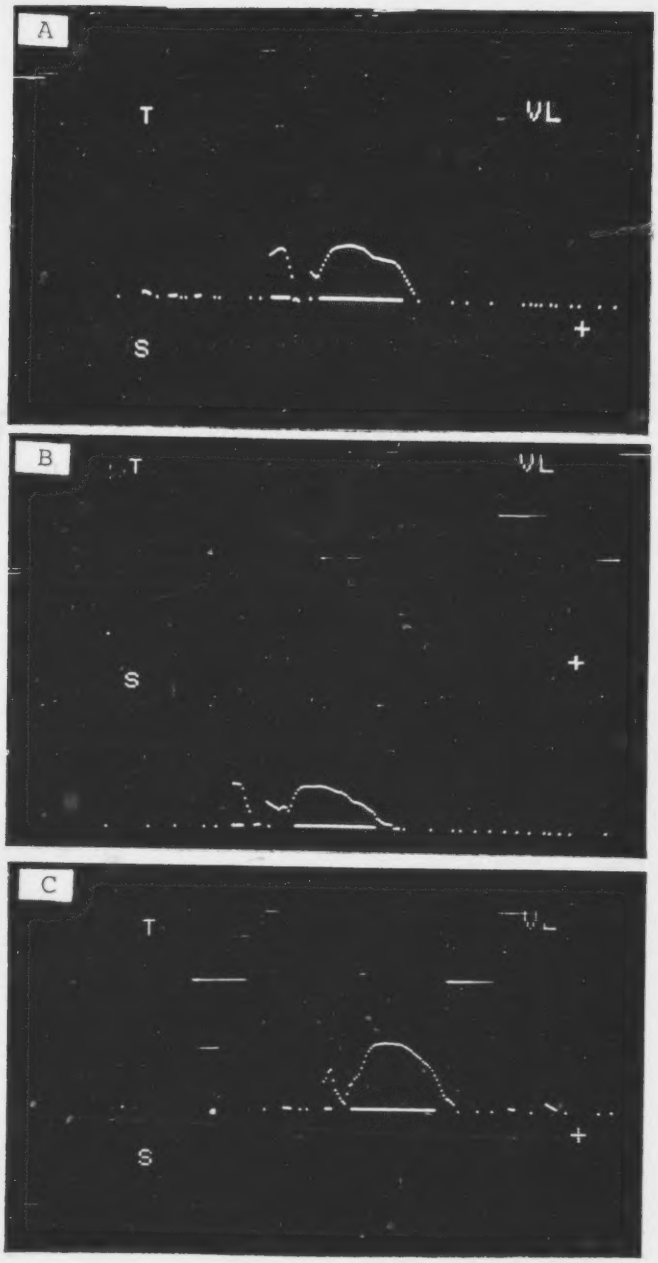


Figure 4. VL display for the utterance a car produced by three different normally-hearing speakers. The final stressed syllable is much longer than the initial unstressed syllable.

shorter than a single monosyllable. The amplitude of the stressed syllable (as indicated by the loudness contour) should usually be greater than that of the unstressed syllable. Examples of the normal timing pattern for the word puppy are shown in Fig. 5.

It is common for deaf students to tend to put equal stress on both syllables of these two-syllable utterances. Thus, the unstressed syllables are usually not sufficiently shortened relative to the stressed syllables. Examples are shown in Fig. 6 for final stress and Fig. 7 for initial stress. Another error that is frequently made is to insert too long a pause between the two syllables, as in Figs. 6b and 7a. Work with the students should be aimed at correcting these two problems. Figure 7c is another example illustrating the gradual drop in loudness in the final syllable, leading to a breathy termination. The utterances with final stress should be regarded in some sense as monosyllabic words with a brief initial syllable appended, rather than a concatenation of two syllables that have equal status. An utterance with initial stress should have about the same total duration as a monosyllable, but it is separated into two parts by one or more consonants.

If a student is having difficulty in producing these two-syllable temporal patterns, it may be convenient to practice with nonsense utterances such as pa pá, or pá pa, so that problems with articulation do not stand in the way of achieving the appropriate timing. When the proper rhythm is mastered with this nonsense material, then these patterns should be practiced with meaningful two-syllable sequences involving a variety of vowels and consonants. The student should be encouraged to maintain reasonable vowel and consonant articulation while producing acceptable timing patterns.

The diagnosis and training of timing for longer utterances involves concepts similar to those introduced in moving from

BEST COPY AVAILABLE

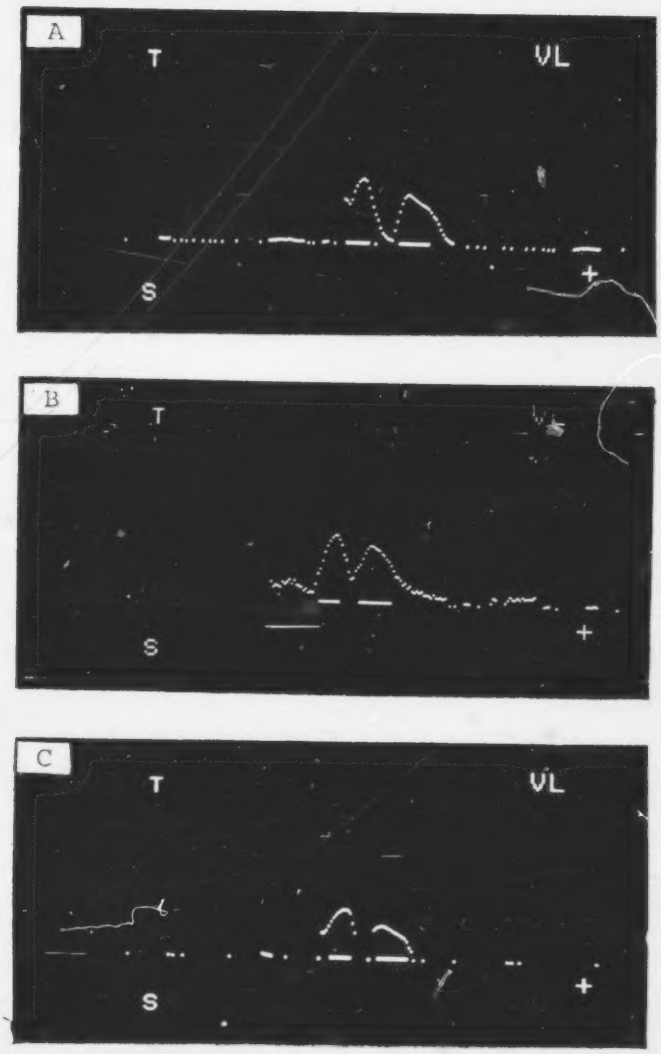


Figure 5. VL display for the word puppy produced by three different normally-hearing speakers. The two syllables are comparable in length for these examples.

BEST COPY AVAILABLE

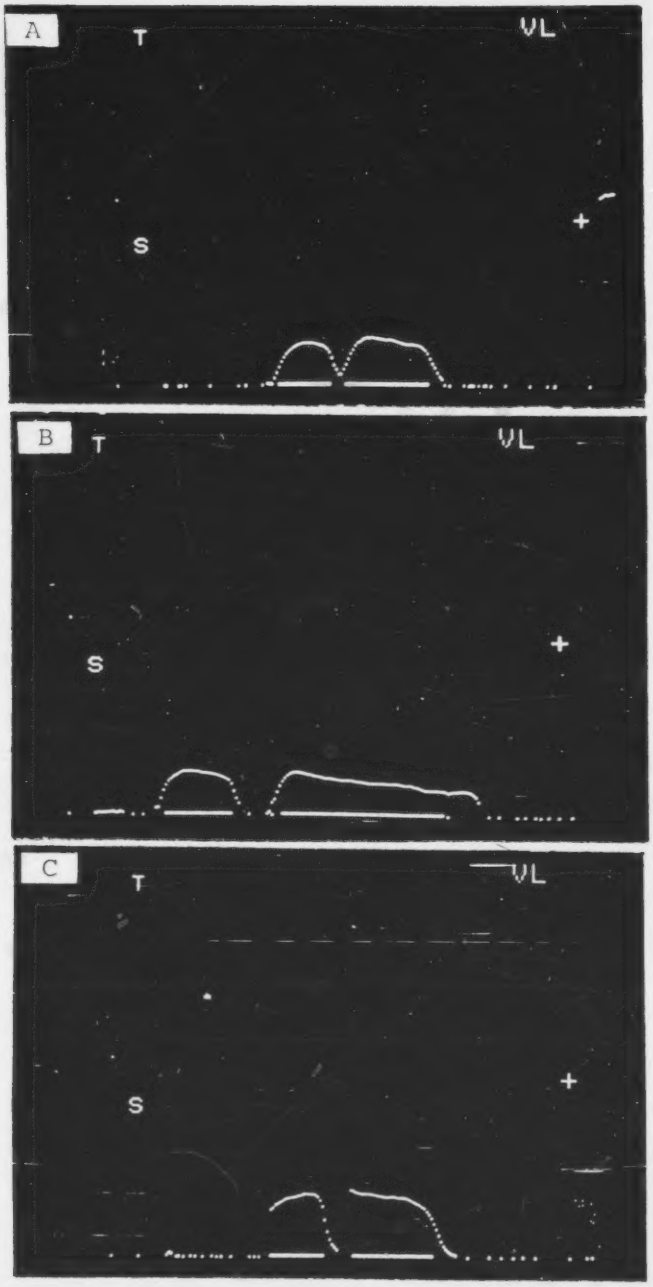


Figure 6. VL display for the utterance a car produced in an unacceptable fashion by three different deaf students.

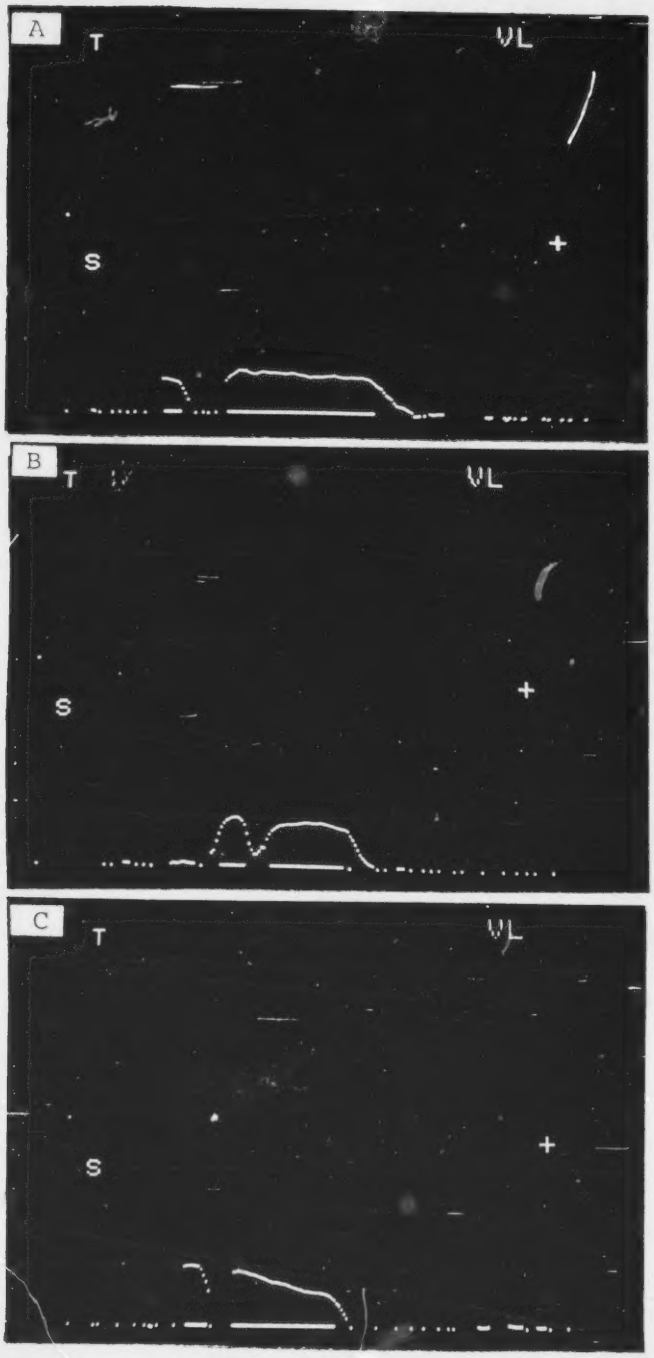


Figure 7. VL display for the word puppy produced in an unacceptable fashion by three different deaf students.

monosyllables to two-syllable utterances. For example, one can append an additional unstressed syllable at the beginning of the utterance the park to produce the phrase at the park. The first syllable should have a duration similar to (perhaps slightly longer than) that of the second unstressed syllable, and both unstressed vowels should be considerably shorter than the final stressed vowel. Examples of the timing pattern for normally-hearing speakers producing this phrase are given in Fig. 8. Voicing for the very brief second syllable sometimes does not register on the display. Again, the tendency of some deaf students is to produce each syllable with approximately the same duration, as illustrated in Fig. 9 (especially Fig. 9c). Training for these students should concentrate on producing a shorter syllable in the unstressed position, with an amplitude that is lower than that of the stressed syllable. Initial work with these stress patterns may be facilitated by the use of nonsense utterances such as pa pa pá. Figures 9a and 9b also illustrate another common problem: the introduction of an inadvertent syllable between two abutting consonants, in this case between at and the.

An example of a different three-syllable stress pattern appears in Fig. 10, for the phrase the father. In this case, the final word should have a temporal pattern similar to that of other two-syllable utterances with initial stress, and the initial unstressed syllable should be appended in much the same way as before, as shown in Fig. 10a for a normally-hearing speaker. The utterance by the deaf student (Fig. 10b) has much too long a final syllable.

As one increases the number of syllables beyond three, the number of possible patterns greatly increases. The same general timing principles apply, however, and the diagnosis and training of timing should focus on these principles. These include: (1) unstressed syllables should be shortened relative to stressed

BEST COPY AVAILABLE

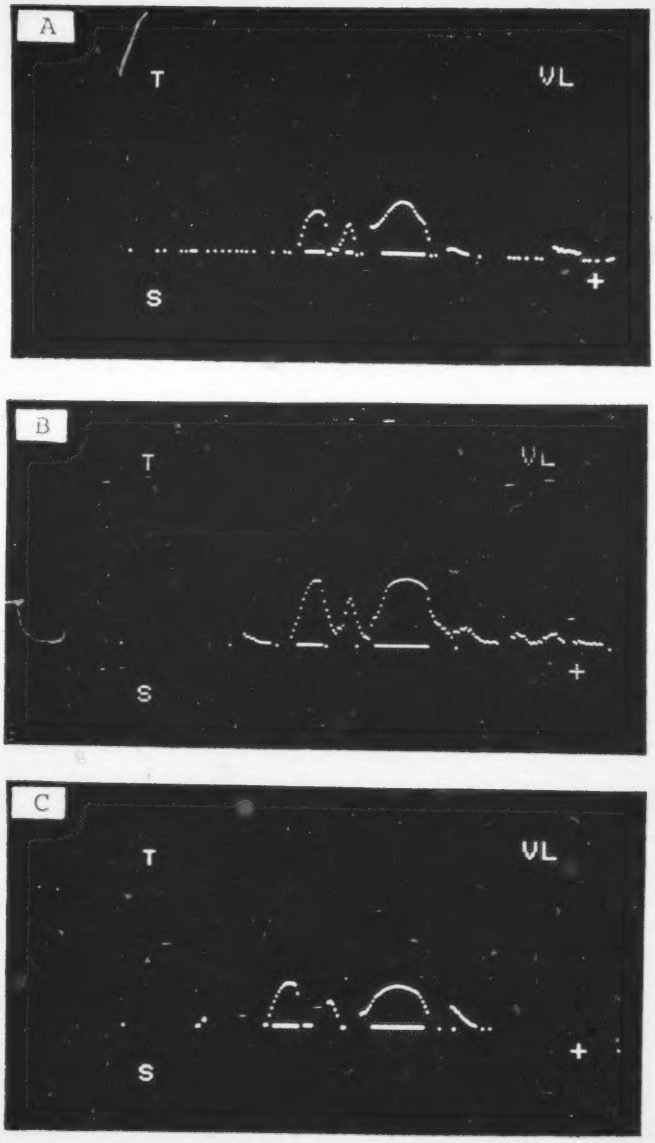


Figure 8. VL display for the phrase at the park produced by three different normally-hearing speakers.

BEST COPY AVAILABLE

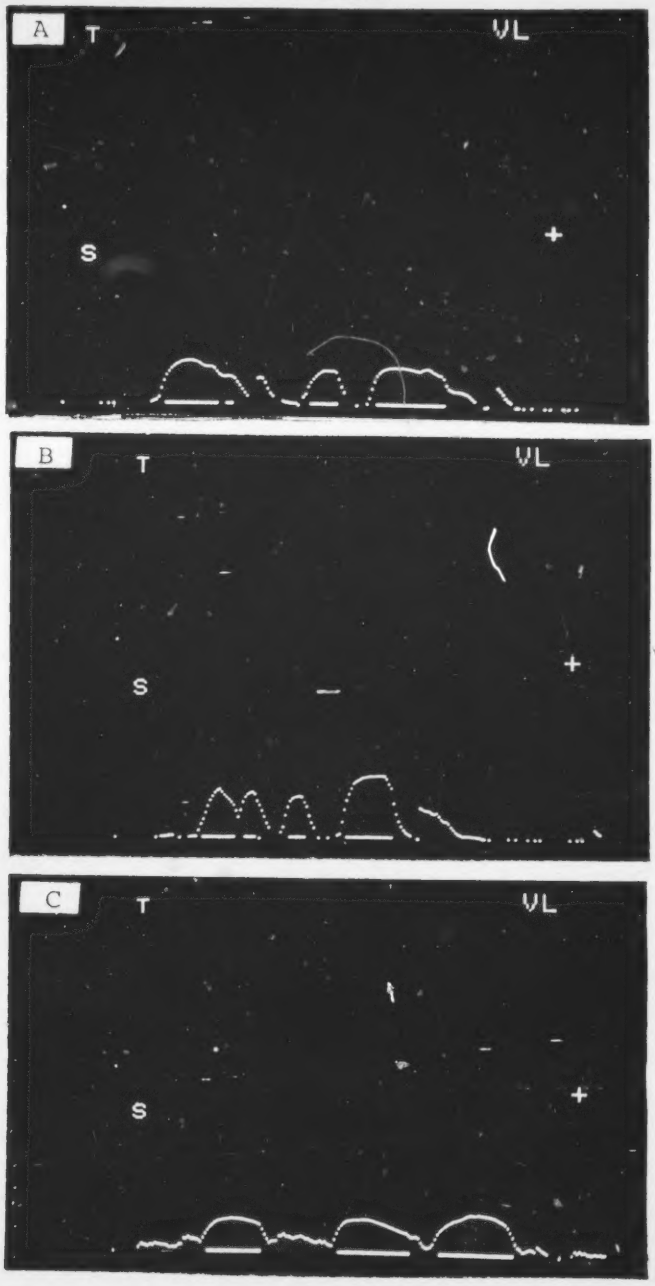


Figure 9. VL display for the phrase at the park, produced in an unacceptable fashion by three different deaf students.

BEST COPY AVAILABLE

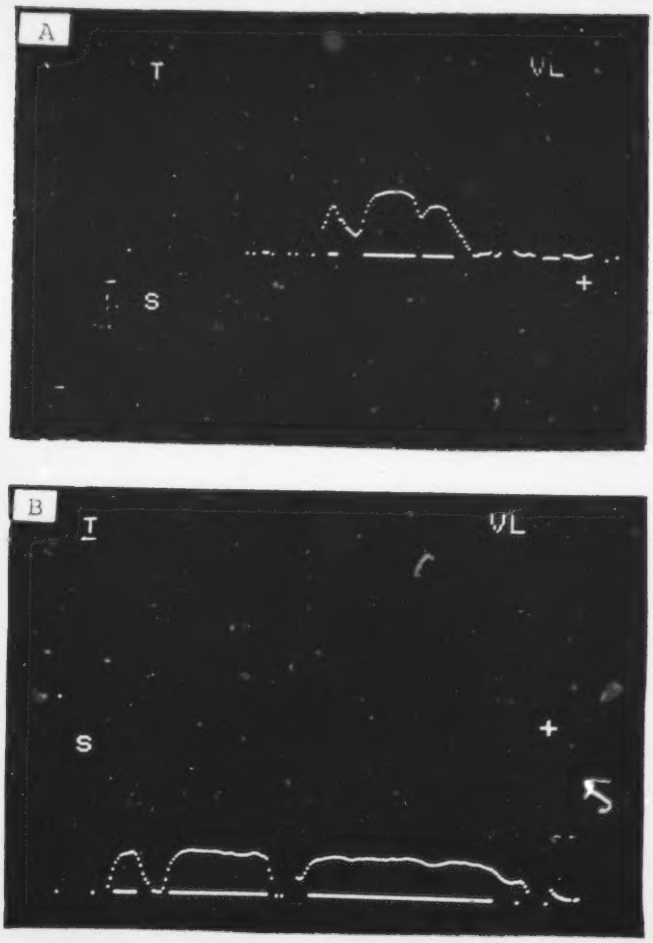


Figure 10. VL display for the phrase the father, produced by a normally-hearing speaker (top) and produced with unacceptable timing by a deaf student (bottom).

syllables; (2) syllables at the end of an utterance, or syllables occurring before a pause, should be lengthened relative to syllables that occur within an utterance; (3) inadvertent pauses within a phrase should be avoided. The second of these rules can result in an utterance-final unstressed syllable that is comparable in length to a preceding stressed syllable, as noted earlier. Rules (2) and (3) introduce the idea of a pause and a phrase. Longer utterances can be divided into phrases, and these phrases can (optionally) be separated by pauses. Thus in the sentence "My sister has a fish," a phrase boundary occurs after sister. The speaker has the option of inserting a pause at this point, in which case sister occurs before the pause, and the final syllable is lengthened. A gap at any other point in the sentence would constitute an inappropriate pause.

The above discussion indicates that the timing of the syllables in an utterance is closely tied to the stress pattern and to the syntactic composition of the utterance. The contour of fundamental frequency (F_0) is also closely related to these aspects of a phrase or sentence. In fact, it is probable that the timing and the F_0 contour are organized in such a way that the appropriate F_0 contour can be accommodated within the temporal pattern. Thus the assessment and the training of timing and of pitch control for deaf students cannot be separated. As a deaf student acquires skill in producing simple timing patterns, it could be appropriate during training to switch to a display of pitch, and to train the student to produce an appropriate pitch contour that is consistent with the timing objective. Comments on the diagnosis and training of pitch control are given in the next section.

4.2 Control of Fundamental Frequency (F_0)

The initial step in the assessment of fundamental-frequency (F_0) control for a student is to measure the habitual F_0 and the range of F_0 in a short sentence and in several monosyllabic words containing various vowels and consonants. The pitch display is used, with the accelerometer attached to the throat. The habitual F_0 can be approximated by asking the student to say a phrase or sentence (e.g., "My name..."), and obtaining an estimate of the average F_0 by setting the horizontal line on the time-plot display so that it is straddled by the pitch contour. The range is obtained by finding the maximum and minimum F_0 within the sentence. The habitual F_0 should be compared with the range of average F_0 values found for normally-hearing children in the same age range, as shown in Fig. 1. If the average F_0 for the student deviates appreciably from the normal range, consideration should be given to training that would cause the student to modify his habitual F_0.

The F_0 deviations measured in the manner described should be roughly 2:1 for normally-hearing speakers. If the range is less than about 1.5:1 for a student, some training aimed at broadening the range should be carried out. If the range is significantly greater than 2:1, there is the possibility of inadvertent jumps or other discontinuities in F_0, such as shifts to falsetto or to some other deviant mode of vocal-cord behavior.

At this stage in the diagnosis, the pitch contour should be examined and the student's utterances should be judged for the presence of abnormal jumps or discontinuities in the pitch pattern. Several types of abnormalities are possible. One example (AK), shown in Fig. 11, is a jump in F_0 that occurs within a vowel. In this case, this kind of discontinuity takes place toward the

BEST COPY AVAILABLE

Figure 11. This photograph shows the presence of a sudden pitch rise at the termination of the utterance "Audrey," an abnormality which characterized this student's speech.

end of many of the vowels produced by the student. Even though the discontinuous shift in F_0 is not large, it gives an abnormal quality to the speech. Another example is a jump to falsetto for some vowels, but use of F_0 within the normal range for others. The pitch display has proved effective in the process of training students to avoid producing these abnormal jumps in pitch.

The pitch contour within a phrase or sentence should have smooth rises and falls of F_0 on appropriate vowels within the utterance, depending on the stress pattern and the grammatical structure. For many utterances, there is some room for variation in the shape of the contour, depending on the meaning the speaker is trying to convey. Diagnosis of the ability of the student to execute these pitch changes involves assessment at two levels: (1) Is the student capable of producing the requisite changes in F_0 with the proper physiological gestures? (2) If he is capable of actualizing the F_0 variations, does he insert these changes at appropriate points within an utterance?

Interpretation of the F_0 contours on the pitch display is sometimes complicated by the fact that some consonants can produce shifts or possible irregularities in F_0 at the onset or offset of an adjacent vowel. Irregularities or "noise" in the contour can also be observed for some speakers who have a particular kind of harsh or breathy voice quality. Example of such aberrations can be seen in contours for both normally-hearing and deaf speakers. Illustrations of these phenomena are in Fig. 12a (final syllable), Fig. 12b (final syllable), Fig. 12c (final syllable), Fig. 13b (artifacts in each syllable, and Fig. 13a (end of first syllable). If a contour has excessive "noisiness," there could be a problem with accelerometer attachment or location on the speaker, or the speaker could have an excessively breathy or harsh voice, rendering

BEST COPY AVAILABLE

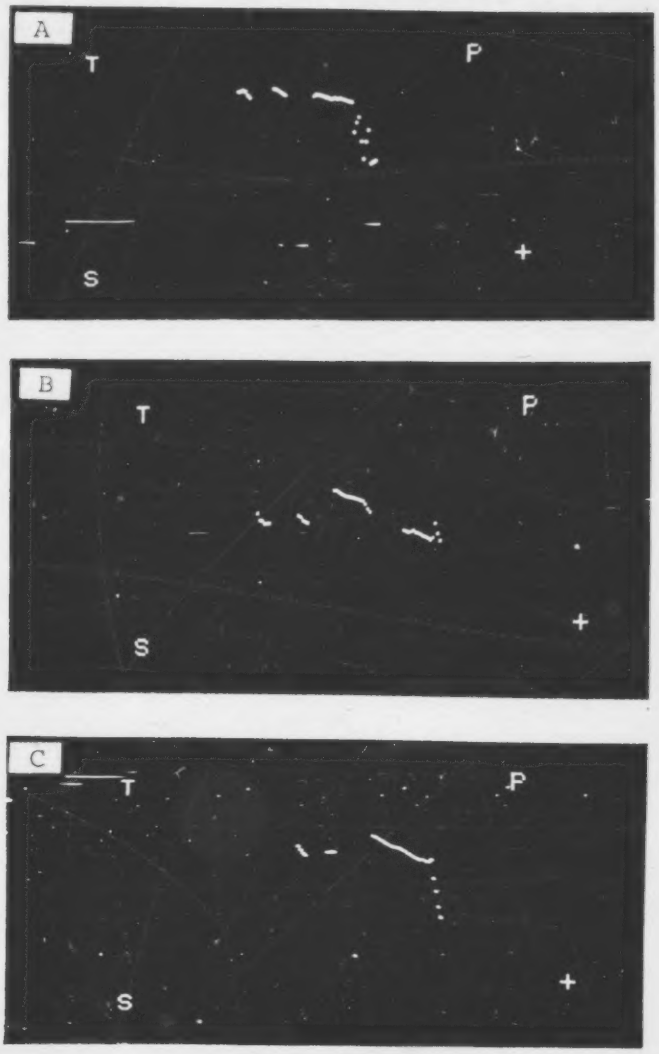


Figure 12. Pitch display for phrases produced by normally-hearing speakers. The top and bottom displays represent the phrase It's a pie; the middle display represents It's a pencil.

BEST COPY AVAILABLE

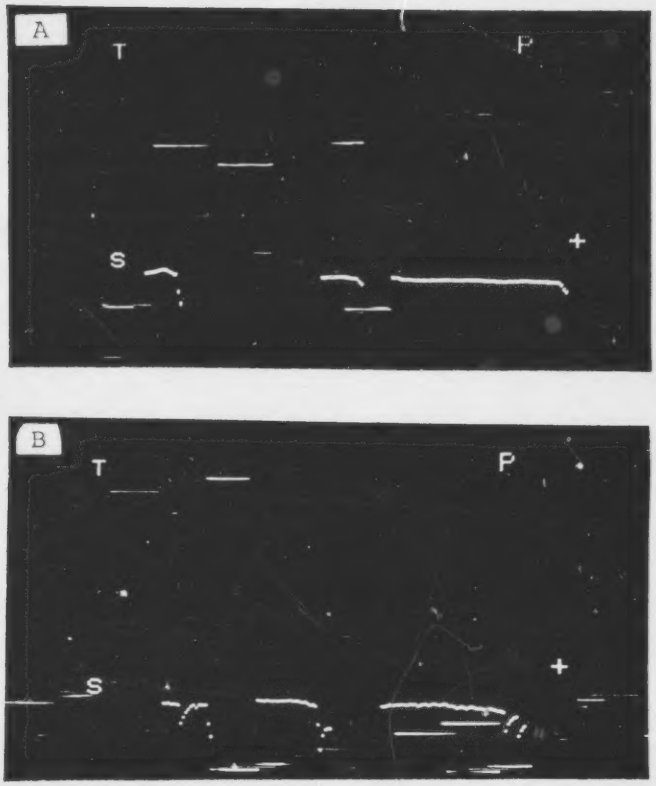


Figure 13. Examples of pitch displays for utterances produced by deaf students with inappropriate pitch contours. The phrases are It's a pie (upper) and the paper (lower).

the measurement of pitch difficult. In assessing the gross changes in pitch to be discussed below, local aberrations in the display of the type just illustrated can usually be ignored.

One of the more basic aspects of a normal pitch contour is the fall in F_0 that must occur at the end of a sentence. If the final vowel in the sentence is stressed, the F_0 fall must occur within the vowel. If the final stressed vowel does not occur at the end of the sentence, the F_0 fall must occur within the unstressed vowel or vowels that follow the stressed vowel.

This terminal fall for utterances produced by normally-hearing speakers is illustrated in Fig. 14. In the phrase "It's a pie," the F_0 fall occurs on the final vowel, which, of course, is lengthened because it occurs before a pause. For each of the phrases "the paper" and "It's a pencil," the final vowel is unstressed, and F_0 on this vowel is lower than on the preceding stressed vowel. Usually the final F_0 fall begins toward the end of the stressed vowel and continues through the following unstressed vowel.

Several kinds of abnormal terminal F_0 contours can be observed in the displays for deaf students in Figs. 15 to 17. One of the most common difficulties is simply failure to produce a significant change in F_0 on or following the final stressed syllable. This problem is illustrated for a final stressed vowel in Figs. 13a, 15a, and 15b, where F_0 remains more or less constant throughout the final vowel. Figure 13b is an example of a level contour for a final unstressed vowel.

When the final syllable should be unstressed, some deaf students inadvertently produce stress on the vowel, Thus, in Fig. 16 there is a terminal fall on the final vowel, but F_0 in the

BEST COPY AVAILABLE

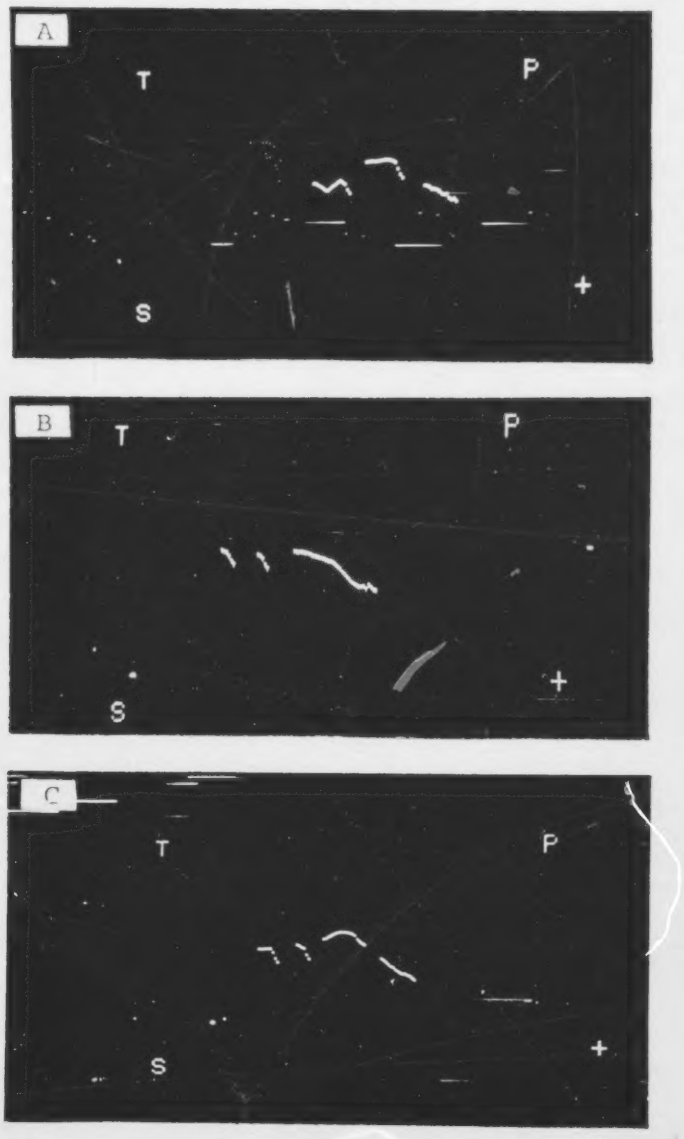


Figure 14. Pitch displays for normally-hearing individuals producing the phrases the paper (top), It's a pie (middle), and It's a pencil (bottom). These contours illustrate the terminal fall in F_0.

BEST COPY AVAILABLE

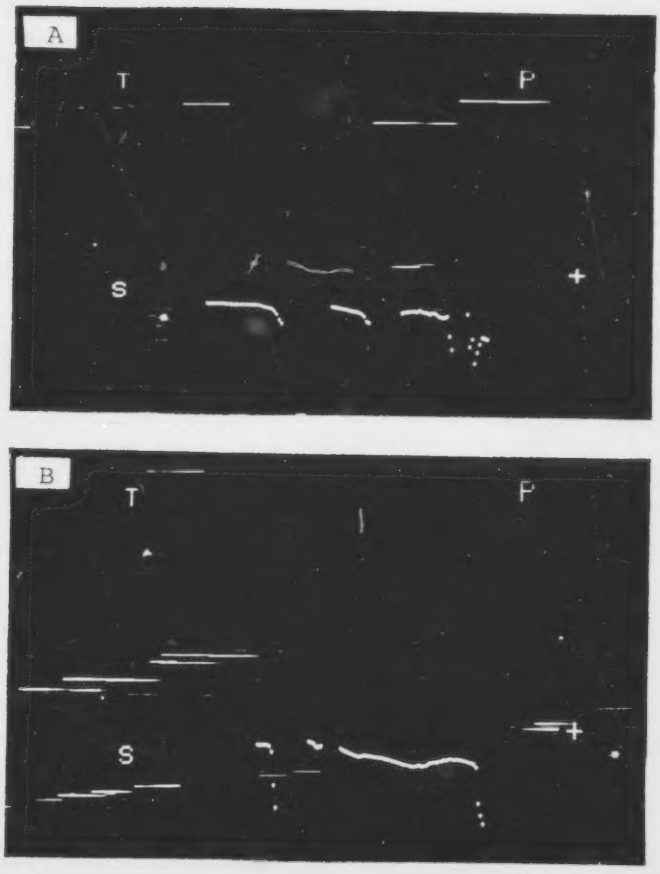


Figure 15. Pitch display for the phrase It's a pie, produced with unacceptable F_0 contour by two deaf students.

BEST COPY AVAILABLE

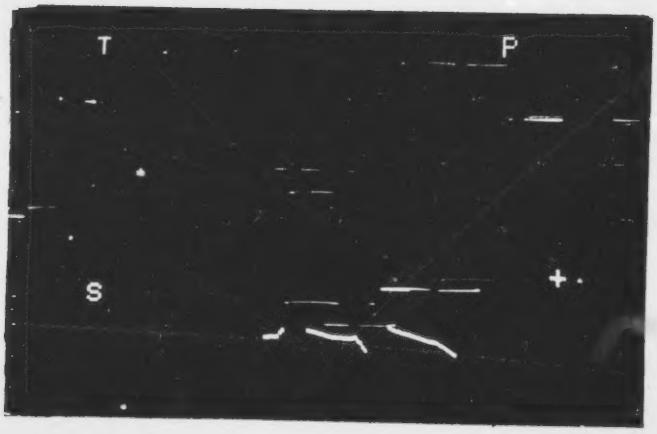


Figure 16. Pitch display for the phrase the paper, produced with unacceptable F_0 contour by a deaf student.

BEST COPY AVAILABLE

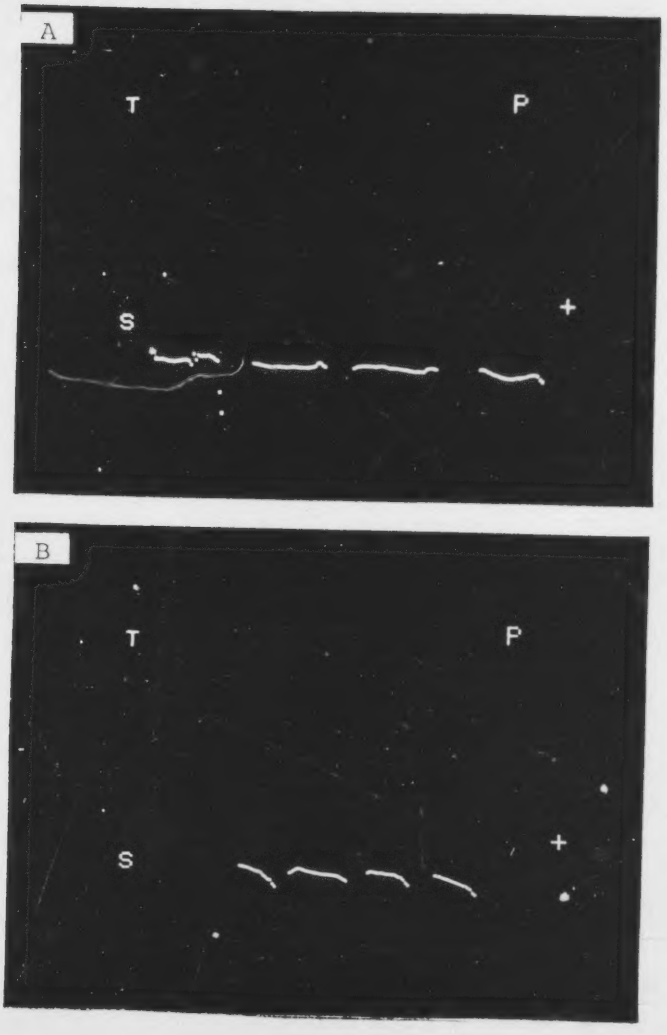


Figure 17. Pitch display for the phrase It's a pencil, produced with unacceptable F_0 contour by two deaf students.

vowel begins above F_0 for the previous vowel. In this example, then, the student apparently has the capability of producing a final fall, but begins this fall on the wrong syllable.

Some deaf students tend to produce each syllable with equal stress, and to place the same pitch contour on each syllable. Examples are given in Fig. 17. In Fig. 17b there is an F_0 fall on each syllable, including the final one, and the problem is the lack of a rise in F_0 on any syllable.

Exercises aimed at training a deaf student to produce a terminal F_0 fall should begin with a simple monosyllabic word or with an isolated vowel. The syllable should be generated with a duration of 500 msec. or less and with an F_0 contour that falls throughout the voiced interval. An F_0 fall of at least 40 Hz (for low-pitched voices, i.e., for older boys) to 80 Hz (for higher-pitch voices) should be set as an objective. A contour that is concave downward is most appropriate (e.g., Fig. 14b), but a brief final levelling of the contour is acceptable, as this example shows. Practice in producing this falling contour on a stressed monosyllable should include syllables with various vowels and with different initial and final consonants.

The next logical step in work with a final F_0 fall is to shift to two-syllable utterances with stress on the first syllable (paper, pencil, baby, etc.). For these words, F_0 should be higher on the first syllable than on the second, and there should be a fall on the second vowel, as in Fig. 14a or 14c. The relative durations of the syllables should be in accordance with the rules for timing discussed in Section 4.1.

Probably the next most basic aspect of the F_0 contour for a sentence is the relatively high F_0 that must occur on the first

stressed syllable in a phrase or sentence. If the first stressed syllable is preceded by one or more unstressed syllables, then F_0 on the unstressed syllables should be lower than that on the stressed vowel. Examples that show a strong initial F_0 rise of this kind are shown in Figs. 12b, 12c and 14a. Such utterances can, however, be produced with a different style, as in Fig. 18a, where the second unstressed syllable is produced with a raised F_0. (In this case, however, the high F_0 on this syllable may be a local influence of the adjacent consonants.)

Speech training aimed at producing a pitch rise from an initial unstressed syllable to a following stressed syllable should utilize initially two-syllable utterances of the type papá, a pie, good-bye, etc. At some point the student should be constrained not only to produce an initial rise to the stressed vowel but also to generate an appropriate final pitch fall as discussed above. The fundamental frequency for the initial unstressed syllable should be greater than the terminal F_0; the highest F_0 should, of course, occur on the stressed syllable. Exercises could also include two-syllable words that are voiced throughout, such as helló, a man, police. etc. With such utterances, however, it may be difficult to determine where each syllable occurs from observations of the pitch display alone. The locations of the syllable peaks can be determined by switching to the loudness display.

A logical next step in training would be to work with utterances containing more than two syllables, but with only one of the syllables being stressed. Examples of such utterances have been given above.

The next stage of complication in the training sequence would be to work with utterances that require an F_0 fall followed by a

BEST COPY AVAILABLE

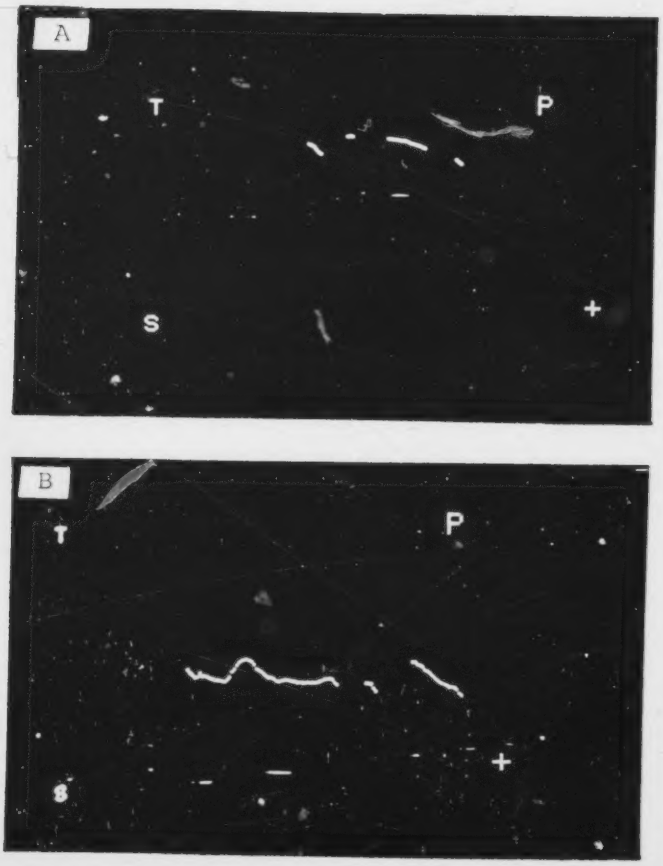


Figure 18. Examples of pitch displays for normally-hearing speakers. (A) The phrase is It's a pencil. The contour shows a rise in F_0 on the second syllable--a somewhat atypical but acceptable way of producing the phrase. (B) The sentence is The boy went to school. The contour shows the rise for the first stressed syllable (boy), followed by a fall, followed by a rise for the word school, with a terminal fall.

rise. Such a sequence of gestures would occur in a sentence with an appropriate syntactic structure containing two stressed syllables separated by one or more unstressed syllables. Examples are: My sister is at home; the boy went to school. In these sentences there is a final F_0 fall on the final stressed syllable, as before, and an initial rise to the first stressed syllable. In addition, most people would produce these sentences with an F_0 fall following the first stressed syllable, an interval of lowered F_0 during the intervening unstressed syllables, and then a rise to the final stressed syllable. An example of such an F_0 pattern is shown in Fig. 18b.

4.3 Velar Control

Diagnosis of the adequacy of velar control consists of assessment of the ability of the student to maintain the velum in a raised position for nonnasal vowels and consonants, to lower the velum for nasal consonants, and to exercise appropriately timed dynamic control of the velum during sequences that involve both nasal and nonnasal sounds.

An initial step is to examine the nasality, using the voicing nasality (VN) display, for monosyllabic words. This assessment is accomplished by first obtaining a reading of nasalization on the display when the student is sustaining a nasal consonant such as m at his normal voice effort. This reading should be in the range 200-300 for most individuals. When a nonnasal word is produced, the peak nasalization should be less than the reading for the nasal consonant by an amount that depends to some extent on the vowel. Minimum difference values that should be achieved for several vowels if they are to have an acceptable degree of nasalization are given in Table 1 (see Stevens, Nickerson, Boothroyd, & Rollins, 1974).

Table 1. Difference in nasalization reading (VN display) for various nonnasal vowels produced in isolated monosyllabic words. These differences are obtained for normally-hearing speakers.

Vowel	Difference in Nasalization Reading*
i	110
a	200
æ	200
u	150

*These numbers are in units of approximately 1/10 dB.

There may be some individuals who produce vowels with an acceptable degree of nasality (as judged by listeners) when the difference values are somewhat less than those in Table 1. Thus the values in Table 1 should be used as guidelines that may be modified up or down as much as 50 units, based on the subjective impression of the teachers.

In many situations, particularly for low vowels (e.g., /a/ and /æ/), the nasalization reading in a nonnasal word is, in effect, zero, and the VN display for the word is simply a horizontal line indicating the voicing interval (as in Fig. 19a). In other cases, the nasalization trace is visible above the voicing line (Figs. 19b and 19c), sometimes for only a portion of the vowel. When an open syllable is produced in utterance-final position, a hearing speaker often lowers the velum toward the end of the vowel in a sort of relaxation gesture that terminates a breath group. Thus, a slight rise in the nasalization reading at the end of a final vowel should not be regarded as an abnormality that requires correction. An example is shown in Fig. 19c.

For the examples given in Figs. 20 and 21, the peak nasality within the word exceeded the criterion given in Table 1 by a substantial amount. The word pad begins and ends with an obstruent consonant (a consonant that requires buildup of pressure in the mouth). Pressure buildup in the mouth requires, of course, a closed velum. In the examples in Figs. 20 and 21, there is evidence in the display to indicate that velar closure was achieved during the consonants, since the nasalization trace was low in the voiced portions adjacent to the consonants. Shortly after the consonant release, however, velar opening occurred, and there was excessive nasalization throughout much of the vowel. These students were apparently able to raise the velum during the consonants, and may have managed to maintain a raised velum through forces resulting from

BEST COPY AVAILABLE

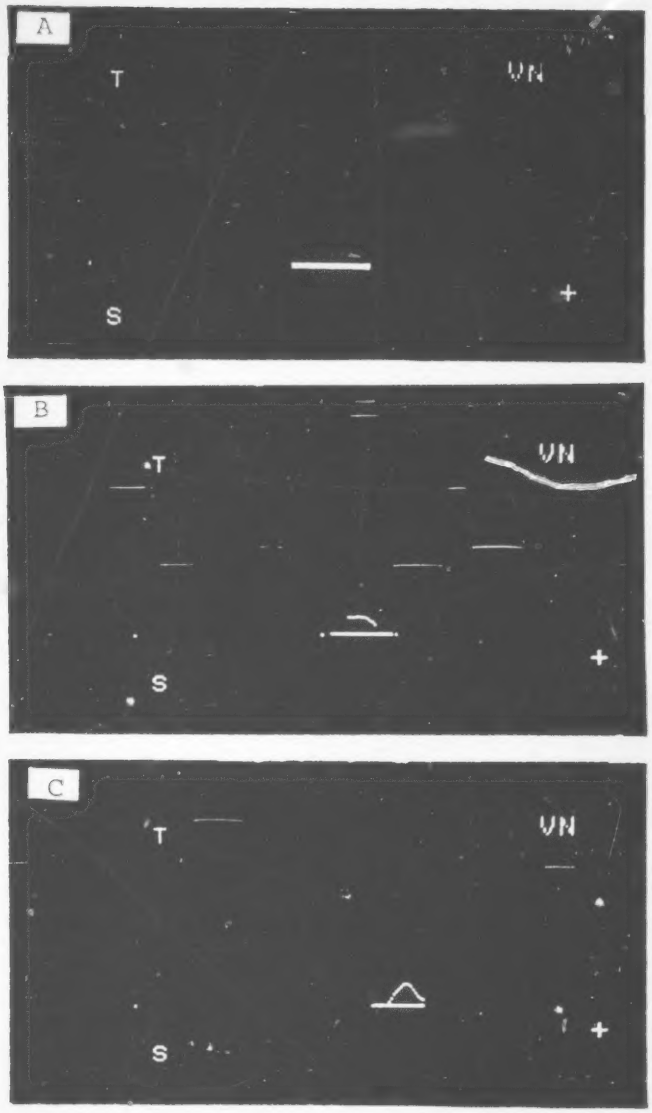


Figure 19. Voicing-nasalization (VN) display for monosyllabic words produced with acceptable velar control by normally-hearing speakers. For the top trace (word pad), the nasalization trace does not show above the baseline, whereas for the middle (pad) and bottom (fee) displays, the nasalization trace can be observed, but it is below the criterion given in Table 1.

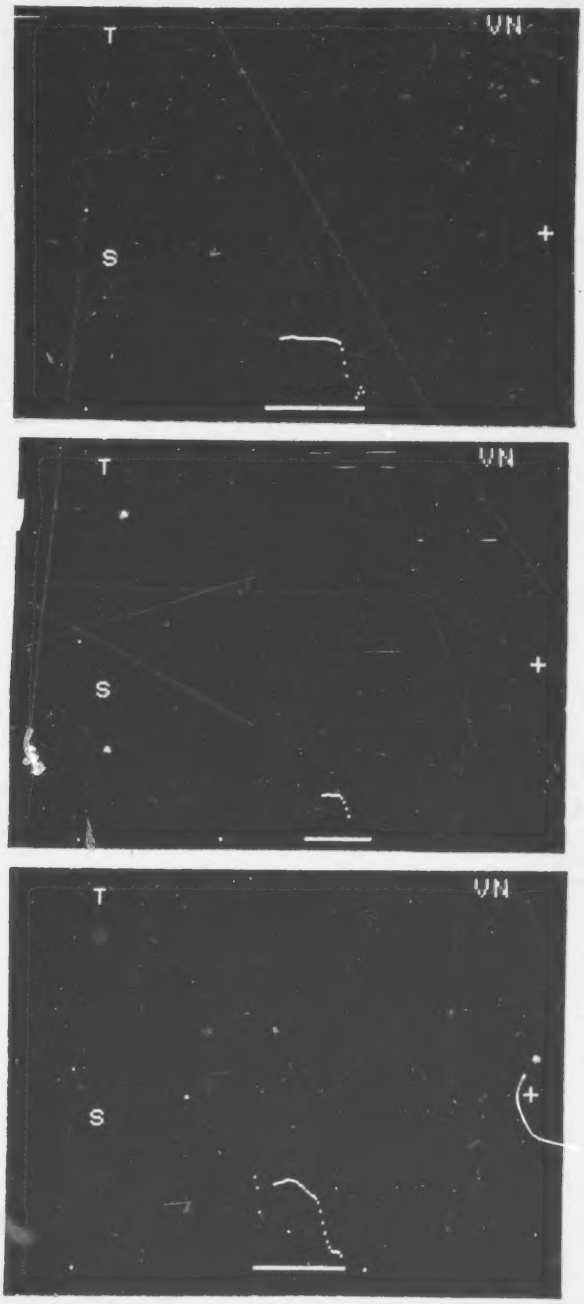


Figure 20. VN display for the word pad, produced with unacceptable nasalization by three deaf students.

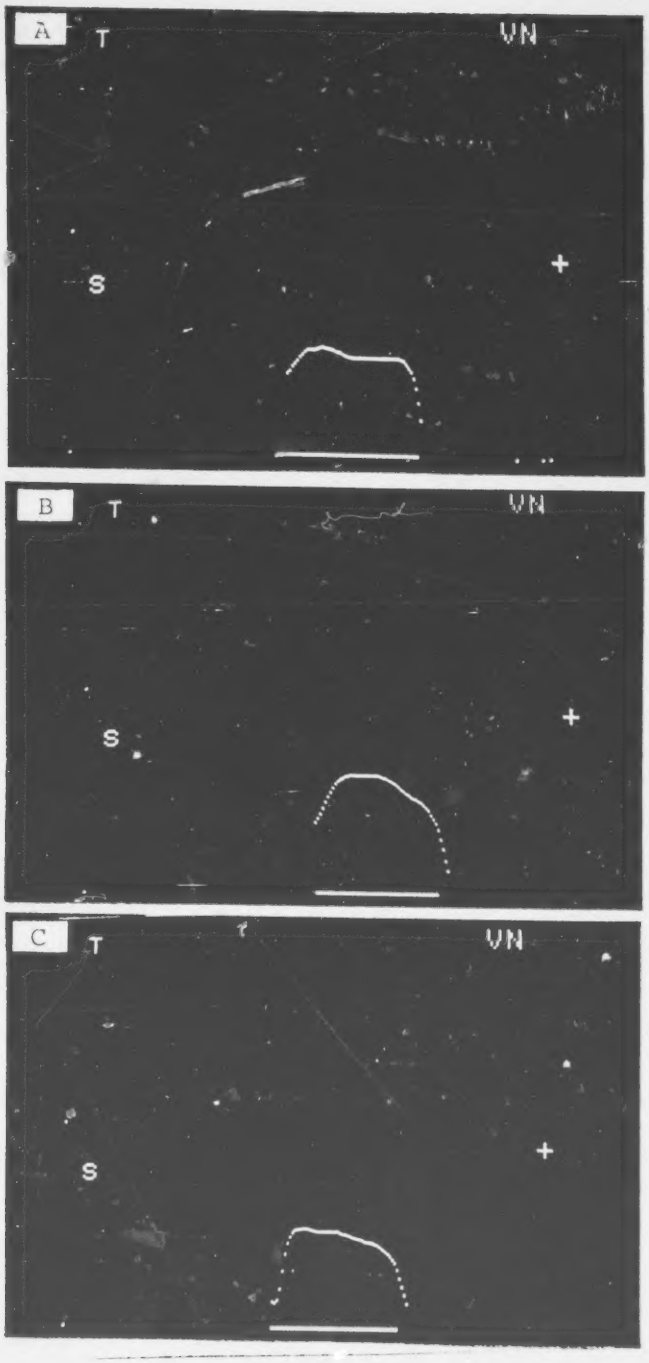


Figure 21. VN display for the word fee, produced with unacceptable nasalization by three deaf students.

the increased pressure in the mouth. A raised velum could not be maintained, however, during the vowel, when there was no increased intraoral pressure.

The initial diagnosis with monosyllabic words should be carried out with words containing several different vowels, such as those listed in Table 1. It frequently happens that deaf students can produce some vowels with an acceptably low degree of nasalization, but not others.

For students who are producing monosyllabic words with excessive nasalization, the initial training should be directed toward achieving control of the velum during vowels. Since the velar position cannot easily be "felt" by the student, some experimentation on the part of the student is needed, with the display operating in real time. Several different strategies can be tried--the strategy that proves to be successful may differ from one student to another.

One strategy is to work initially with isolated vowels, attempting tc reduce the nasalization reading during the course of a long vowel of several seconds' duration. Different vowels can be tried. If success is achieved with one vowel, the student should attempt a shift to another vowel while maintaining a closed velum. Goals can be set for the student--a certain number of successes at a particular task being required before proceeding to the next.

Another strategy is to work with monosyllables rather than with sustained vowels. If, for example, the student can raise the velum to produce an initial stop consonant such as p, he should then try to maintain the low nasalization reading throughout the vowel. Exercises with a sequence of syllables such as pa pa pa may be helpful in learning this task.

A next step in the diagnosis of adequacy of velar control is to examine monosyllabic words containing nasal consonants. The display for these words should show a substantial peak in nasalization for the nasal consonant, and usually a lesser degree of nasalization in the adjacent vowel. It is not necessary for the nasalization reading in the vowel to satisfy the criteria in Table 1 if the vowel is adjacent to a nasal consonant. Some deaf students may exhibit hyponasality in the utterances containing nasal consonants, as manifested by low nasalization readings within the consonant. That is, the nasal consonant is inadvertently produced with a raised velum, resulting in a stop consonant rather than a nasal. Training materials designed to aid in the correction of this problem should consist initially of monosyllables with nasal consonants in initial position and in final position in the syllable.

The ability of a student to produce velar opening and closing with acceptable timing should be assessed initially with simple utterances containing both nasal consonants and obstruents. The simplest utterances of this type are monosyllabic words such as nap and Sam. The first of these words requires a lowered velum at the onset of the word, and the velum must then be raised at some time during the vowel, in preparation for the final obstruent consonant. A raised velum is, of course, essential if mouth pressure is to be built up in the final consonant. The second word is produced with the reverse sequence of gestures: i.e., a raised velum during the first consonant, and a lowering gesture during the vowel in preparation for the final nasal consonant. Figure 22a shows an example of the normal sequence, with initial low nasalization (immediately following the onset of the vowel), a somewhat nasalized vowel, and a peak in the nasalization contour for the final consonant. Some of the deviations from this prototype pattern that might be expected for normally-hearing listeners are illustrated in Figs. 22b and 22c. In the former case, the peak in nasalization in the

BEST COPY AVAILABLE

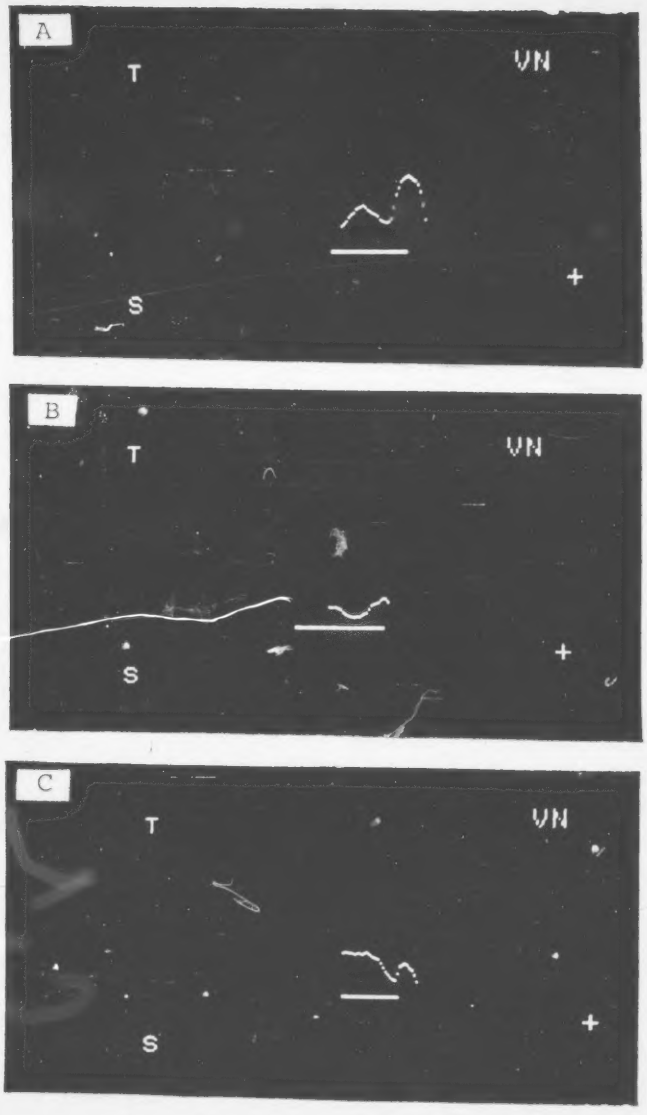


Figure 22. VN display for the word Sam, produced by three normally-hearing speakers. The example in (A) is most typical of such speakers.

consonant is rather small, indicating that the final consonant was not fully nasalized. The latter display shows strong nasalization of the vowel, and an apparent drop in amplitude at the end of the word. (Note that the /s/ sound should not be visible on the VN display.)

Monosyllabic utterances that involve a more complex sequence of velar movements consist of words with consonant clusters containing nasals--words such as snap, Smith, drink, jump. In these words, the velum must undergo a raised-lowered-raised sequence of movements, since the initial and final consonants are both ostruents.

Control of the velum becomes slightly more complex in two-syllable words containing both nasals and obstruent consonants. An example is the word Monday, illustrated in Fig. 23. For the normally-hearing speakers, nasalization peaks are observed for the two nasal consonants in the first syllable, and the intervening vowel is usually produced without velar closure. The velum must close rapidly in preparation for the stop consonant d, and then remain closed during the final syllable, as the display in Fig. 23 shows. The deaf students represented in Fig. 24 produced adequate nasalization in the first syllable, and apparently raised the velum to produce the stop consonant d. However, there was excessive nasalization in the final syllable, which should, of course, be produced with velar closure. In Fig. 25 the student apparently failed to open the velum in the first syllable--an example of hyponasality.

For the word Monday, the velum is lowered during the first syllable, and then must be raised for the second syllable. The reverse sequence must be produced in words such as happen and patting. Both types of words should be included in a training sequence aimed at the development of improved control of the velum.

BEST COPY AVAILABLE

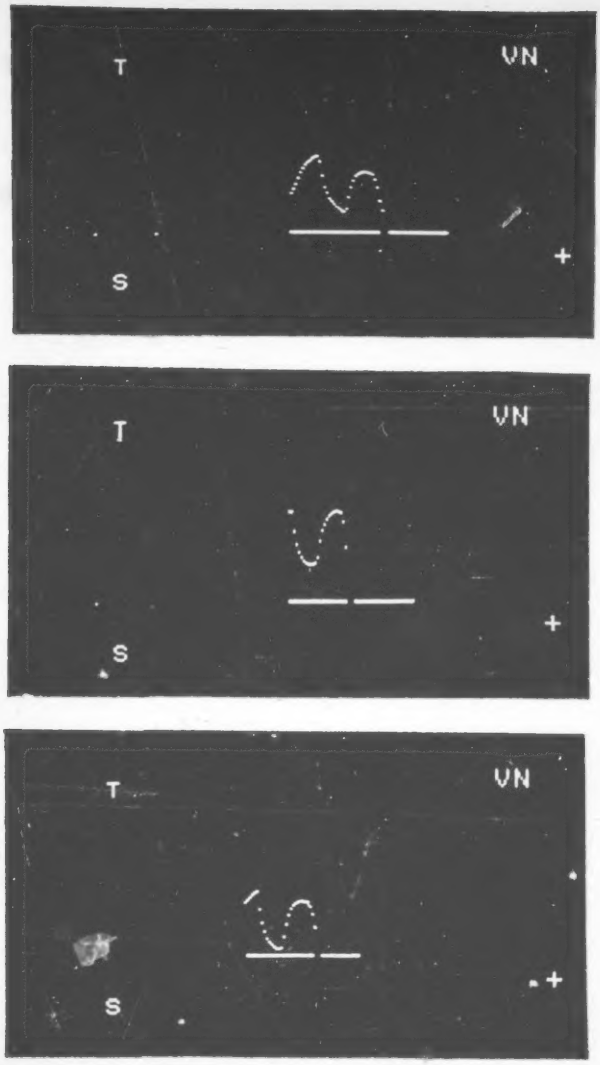


Figure 23. VN display for the word Monday, produced by three normally-hearing speakers.

BEST COPY AVAILABLE

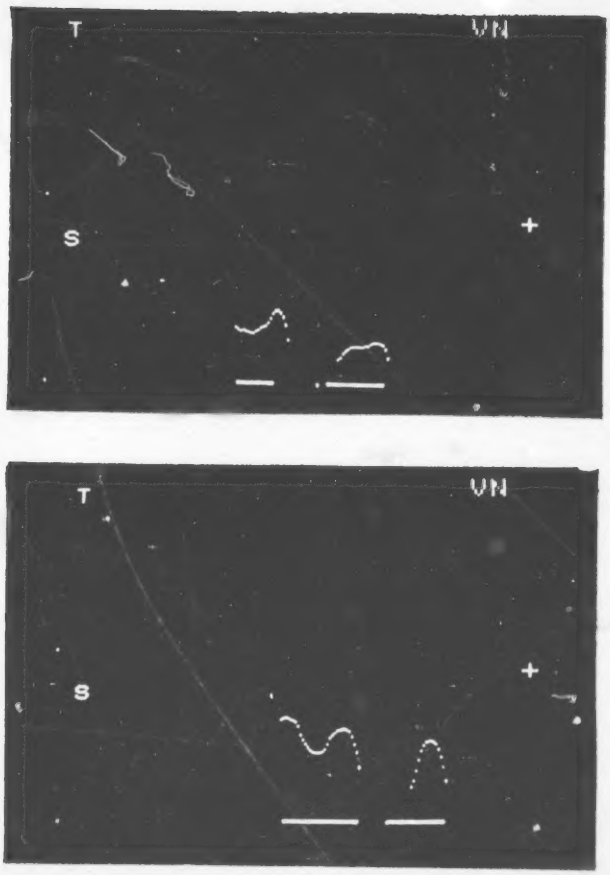


Figure 24. Examples of VN display for deaf students producing the word Monday with excessive nasalization on the second syllable.

BEST COPY AVAILABLE

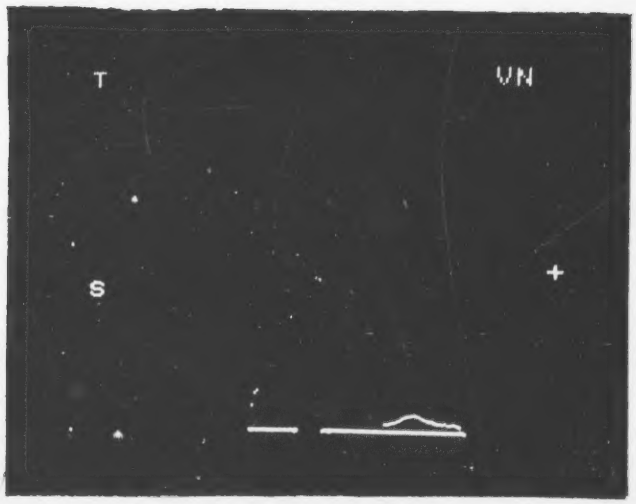


Figure 25. VN display for a deaf student producing the word Monday with hyponasality. The velum apparently remains raised throughout the word.

Diagnosis and training of velar control for longer utterances should include phrases containing no nasals and phrases containing mixed nasal and obstruent consonants. Examples are the short sentences "Be my baby" and "You can drink your milk," for which displays for normally-hearing speakers are shown in Figs. 26 and 27. For the first of these sentences, the vowel in the word be is nasalized because it precedes a nasal consonant, but the two syllables of the last word show no nasalization. The second sentence (Fig. 27) contains three nasal consonants, and only the first word, which is separated from a nasal consonant by a stop consonant, is free of nasalization. The unacceptable samples for the first of these sentences include cases where the velum remains raised during the nasal consonant (Fig. 28a) and where nonnasal vowels are inadvertently nasalized (Figs. 28b and 28c). Similar examples of hyponasality and hypernasality can be seen in Fig. 29. In this sentence, the word you should not be nasalized since it is separated from the nasal consonants by an obstruent, for which the velum must be raised. Other vowels, such as the vowel in your, may become nasalized since the velum is lowered in anticipation of the nasal consonant in the following word. It is not uncommon for a given student to denasalize some nasal consonants and to inadvertently nasalize some vowels within the same utterance--i.e., the inability to control the velum with the proper timing results in both hyponasality and hypernasality. An example is given in Fig. 29a, where the first vowel is inadvertently nasalized, and the /m/ in milk is produced as a stop consonant.

4.4 Articulation

The training of deaf children to achieve improved articulation involves work with many different types of articulatory gestures. The displays that have been implemented to date are not sufficiently versatile to be used to indicate a wide variety of articulation problems. This inadequacy is due in part to our lack of knowledge

BEST COPY AVAILABLE

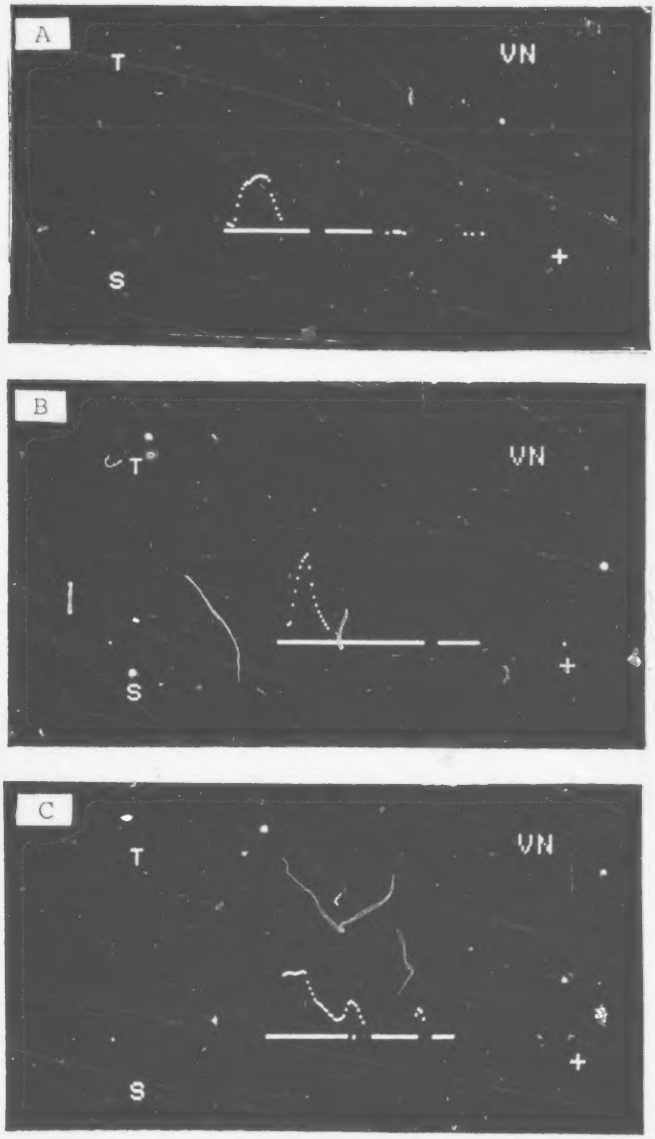


Figure 26. VN display for the sentence "Be my baby" produced by three normally-hearing speakers.

BEST COPY AVAILABLE

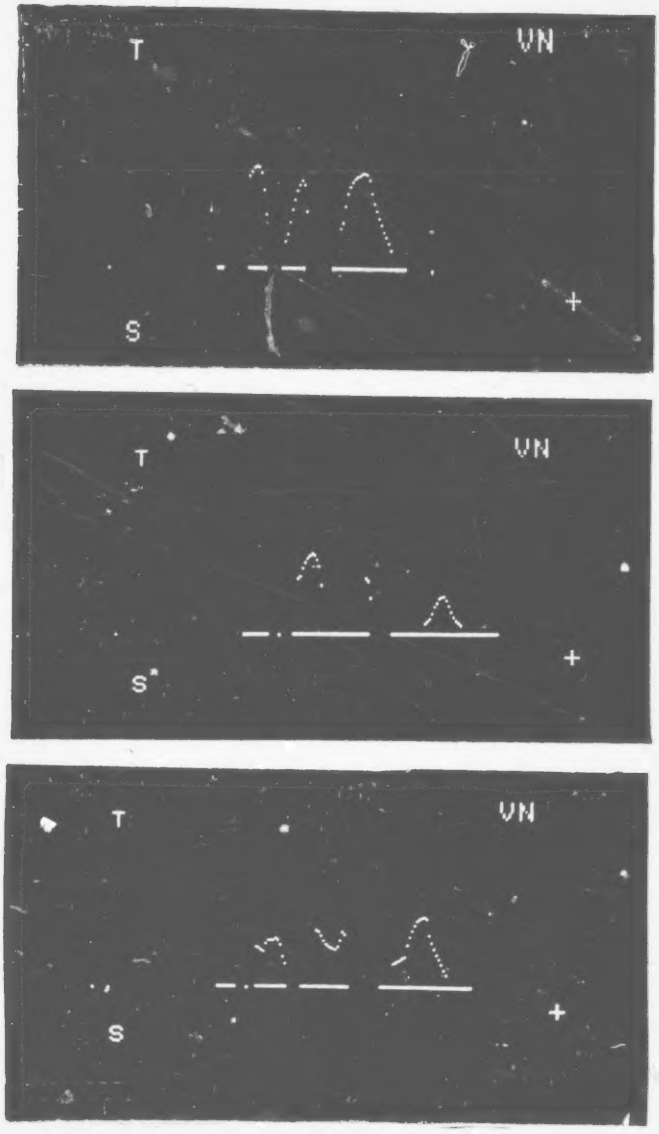


Figure 27. VN display for the sentence "You can drink your milk" produced by three normally-hearing speakers.

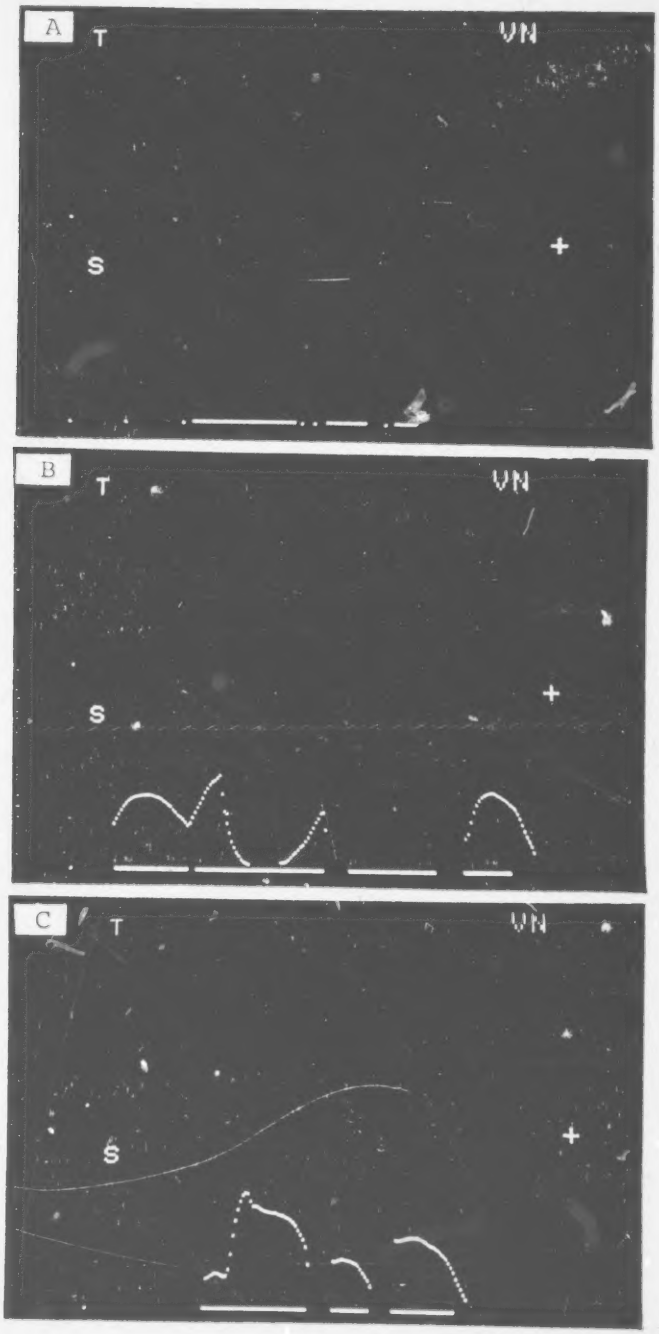


Figure 28. Examples of VN display for deaf students producing the sentence "Be my baby" with improper control of the velum.

BEST COPY AVAILABLE

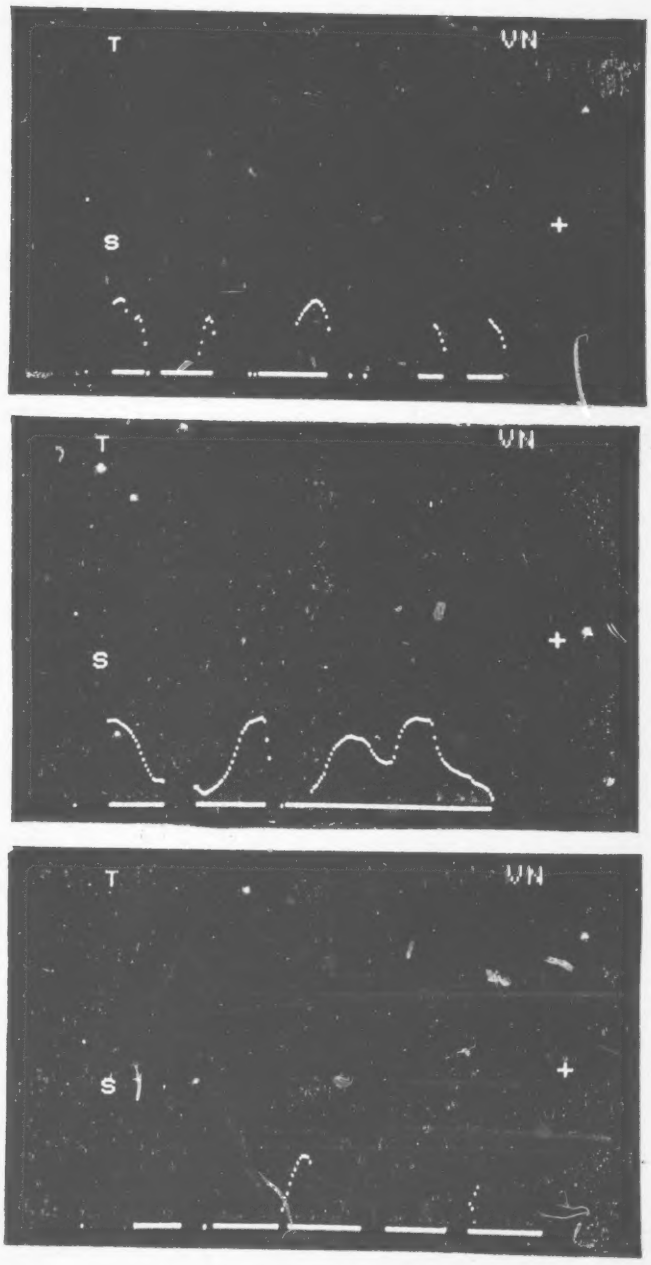


Figure 29. Examples of VN display for deaf students producing the sentence "You can drink your milk" with improper control of the volum.

as to what acoustic parameters can be extracted to give reliable indications of articulation problems and in part to difficulties with providing a display that can be properly interpreted by an observer.

The dual spectrum display is useful for assessing the adequacy of sounds that can be sustained and can be produced in isolation. It can also be used (with somewhat reduced facility) for displaying such sounds in brief monosyllabic words and for examining sounds that are characterized by a relatively slowly changing spectrum, such as diphthongs. The slow-motion replay option provides the capability for examining the spectral changes in a captured speech sample after the fact, and can be helpful in identifying missing, spurious, or misarticulated sounds.

All of the spectrum displays shown in this section can be generated with a normalization feature that sets a maximum width for the display when the intensity of the sound exceeds a certain threshold. For a sound intensity above this threshold, the shape of the pattern for a given sound should be relatively independent of intensity. When the intensity is below this threshold, the width at all points along the height of the pattern is reduced by the same amount. This can alter the appearance of the pattern, particularly when the width becomes very narrow or reduces to a single vertical line in some regions of the pattern.

4.4.1 Fricative Consonants

The vertical spectrum display gives rather distinctive patterns for the fricative consonants /s/ and /š/. Initial work with these consonants should involve producing them in isolation with the appropriate shape for the display.

Examples of patterns of /s/ produced by normally-hearing speakers are shown in Fig. 30. The distinguishing feature of this pattern is a wide top, narrowing down rapidly to a rather thin "stem." The usual type of inadequacy in the production of /s/ by deaf students is for the pattern to be too wide in the stem immediately below the wide top. An example is shown in Fig. 31a. For some unacceptable examples, there may be no abrupt widening at the top of the pattern, as in Fig. 31b.

Deaf students should be instructed to adjust the position of the tongue tip and to shape the tongue blade while producing a continuous noise, in an attempt to obtain the desired pattern shape. If an adequate continuous sound can be produced, then the next step is to generate a series of brief versions of the sound in isolation, each with the proper shape of the vertical spectrum.

Work should then shift to simple monosyllabic words that begin or end with /s/. Unlike the time-plot displays that are used in the training of pitch, timing, and nasality, the vertical spectrum display does not show a trace of some aspect of an utterance over time. When a fricative consonant is produced in a word, therefore, the spectrum display changes very quickly during the utterance. The display can, however, be replayed and "frozen" at any point throughout a 1-second utterance. In working with monosyllabic words, therefore, the procedure is to store the spectral information for the word (produced by the student or by the teacher) by depressing the appropriate control button, and then to replay the spectrum sequence in slow motion, stopping the replay within the sound that is being worked with--in this case, the /s/. The details of the spectrum can then be examined by teacher and student, and the adequacy of the production can be assessed. A successful version of the student's utterance can, if desired, be shifted to

BEST COPY AVAILABLE

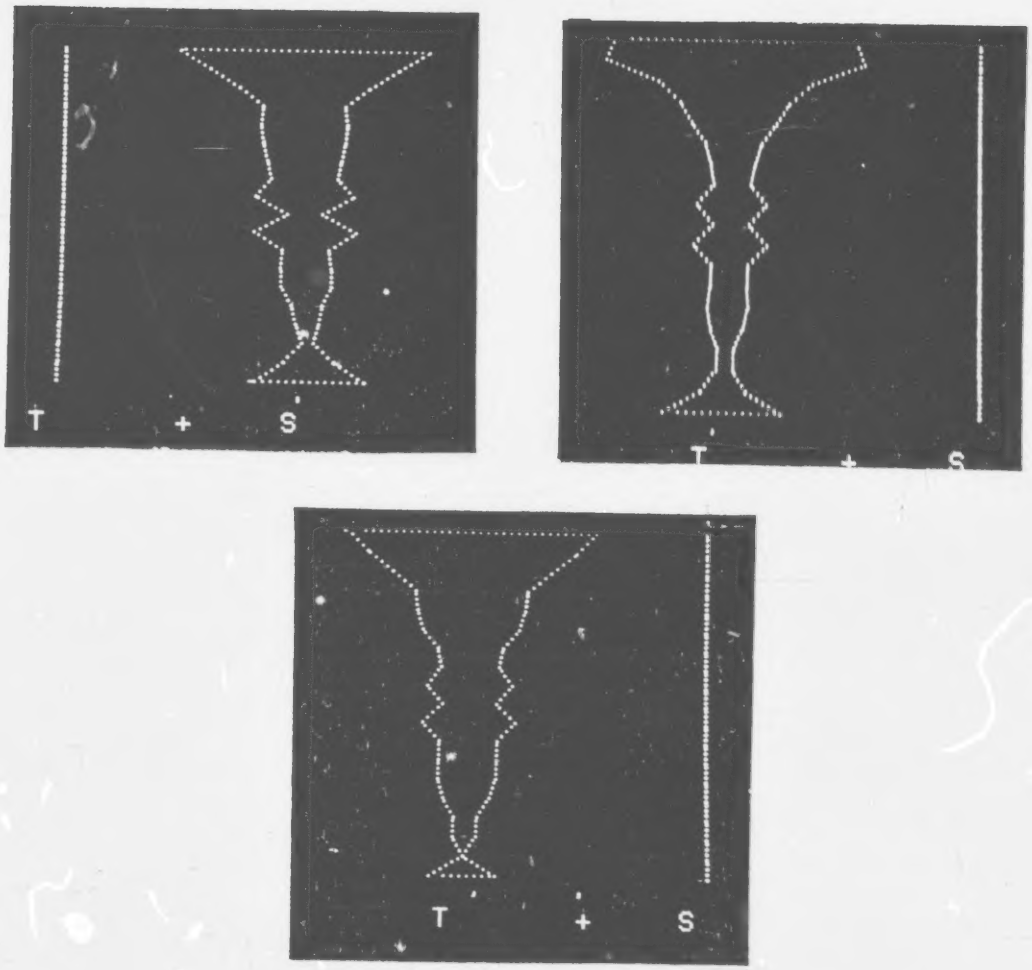


Figure 30. Examples of vertical spectrum display for acceptable versions of the sound /s/.

BEST COPY AVAILABLE

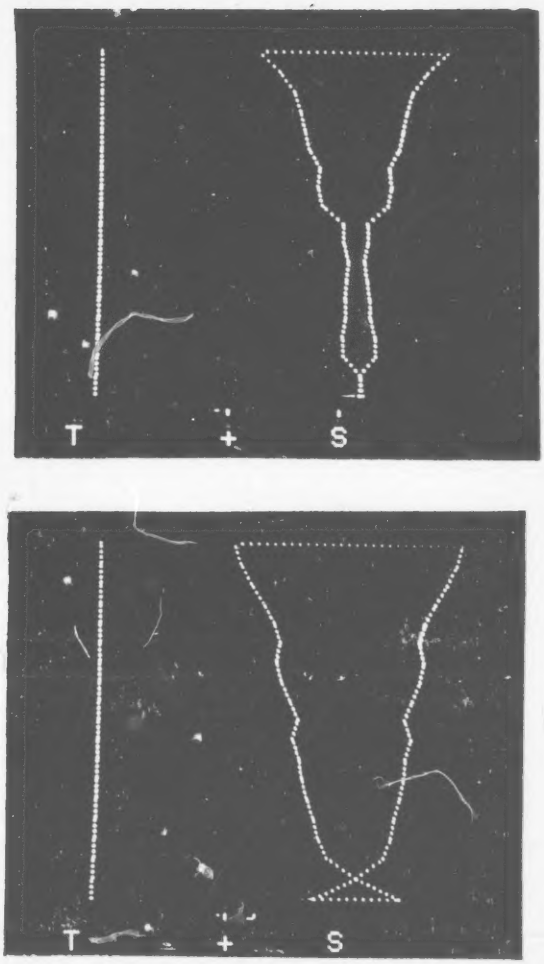


Figure 31. Examples of vertical spectrum display for two deaf children producing unacceptable versions of the sound /s/.

the teacher's side of the display, where it can serve as a target against which future utterances of the student can be compared.

In addition to having the proper spectrum shape, the /s/ in a monosyllabic word should have an appropriate duration in the range 100-150 msec. This means that the s-duration should generally be somewhat less than the duration of the vowel in such a word. The adequacy of the timing of the fricative can be assessed using the VL display, which shows the voiceless fricative as a trace with appreciable loudness but with no voicing.

A next exercise for the deaf student is to produce the /s/ in an intervocalic position such as /asa/ or /isi/, with proper spectrum shape and duration. Work with other phonetic environments should then move to /s/ in initial consonant clusters (e.g., spot, small) and in final consonant clusters (pots, mask). Until an adequate fricative consonant is produced with words such as these, the dual spectrum display should be used. Proper timing of the /s/ production in relation to other aspects of the word should be examined using the VL time-plot display.

A similar sequence of exercises should be followed with /s̬/. Examples of adequate versions of this sound are shown in Fig. 32. The pattern for both /s/ and /s̬/ has a wide top which narrows toward the bottom. The upper half of the pattern for /s̬/ is, however, considerably wider than it is for /s/. A typical difficulty for /s̬/ is a pattern that is too wide in the middle and lower regions, as in Fig. 33.

Extensive experience with training with other classes of fricative consonants, such as the voiceless /f/ and /θ/, the various voiced fricatives, and the affricates, has not yet been obtained.

BEST COPY AVAILABLE

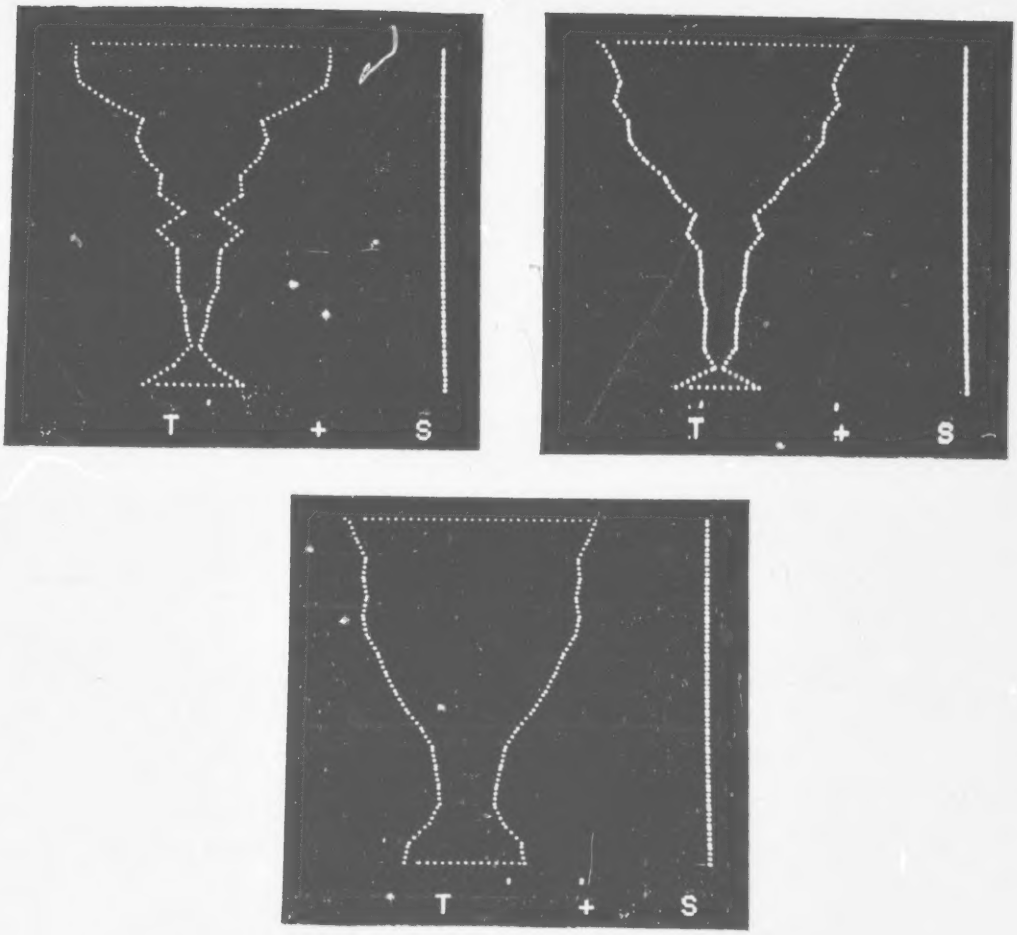


Figure 32. Examples of vertical spectrum display for acceptable versions of the sound /š/.

BEST COPY AVAILABLE

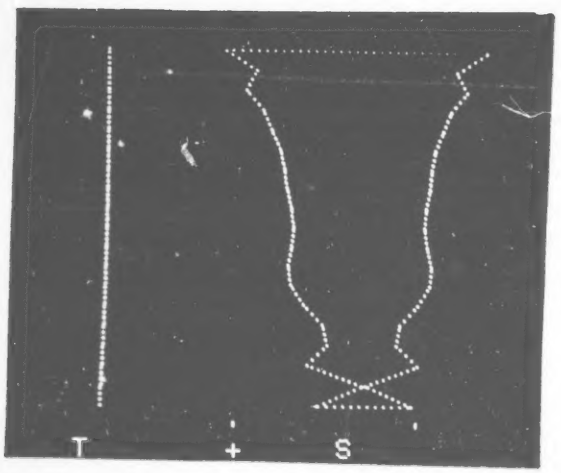


Figure 33. Vertical spectrum display for an unacceptable version of /š/, produced by a deaf student.

4.4.2 Vowels and Diphthongs

The dual spectrum display can indicate the gross acoustic attributes of vowels, and thus can be used to aid in the training of some aspects of vowel production. Examples of this display for the three vowels /a/, /i/, and /u/ produced by normally-hearing speakers are shown in Figs. 34, 35, and 36.

The distinguishing attributes of the pattern for /a/ include: (a) the pattern is pinched in at the bottom; (b) the lower one-third of the pattern is characterized by a wide bulge; (c) the pattern is pinched in somewhat at about the middle or slightly above the middle (Fig. 34c is slightly atypical), but is not too narrow at this point (in comparison with the displays for /i/ and /u/; and (d) the pattern fattens out to produce another bulge in the upper half.

Examples of deviations from this normal pattern for /a/ are shown in Fig. 37. Typical problems include a lack of narrowing of the pattern at the low end (Fig. 37a), two well-separated bumps instead of one wide bump in the lower half or two-thirds of the pattern (Fig. 37b), and a narrowing of the display too far above the middle (Fig. 37c).

In the case of /i/, the pattern begins wide at the bottom and then narrows down to a rather small width within the lower one-third of its height. The region of narrowing may extend over some distance. The pattern then widens in the upper half, so that the width is comparable to, or even greater than, the width at the bottom. The large bump in the upper half may have various detailed fluctuations, as the examples in Fig. 35 indicate. Examples of poor productions of /i/ have patterns with insufficient or even no narrowing in the middle (Figs. 38a and 38c), a narrowing too high

BEST COPY AVAILABLE

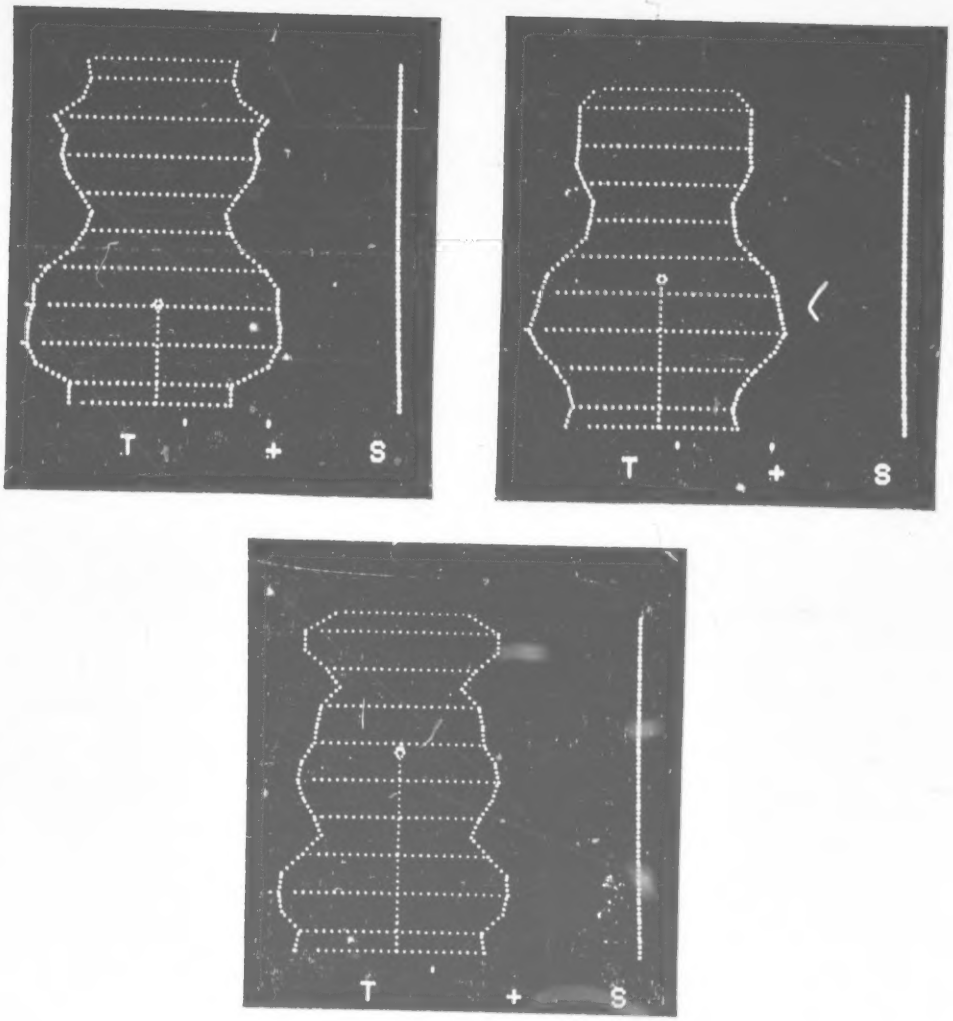


Figure 34. Vertical spectrum display for the vowel /a/, produced by three normally-hearing speakers. The horizontal lines indicate that the sound is voiced. The height of the vertical "lollipop" within the pattern indicates the pitch of the sound.

BEST COPY AVAILABLE

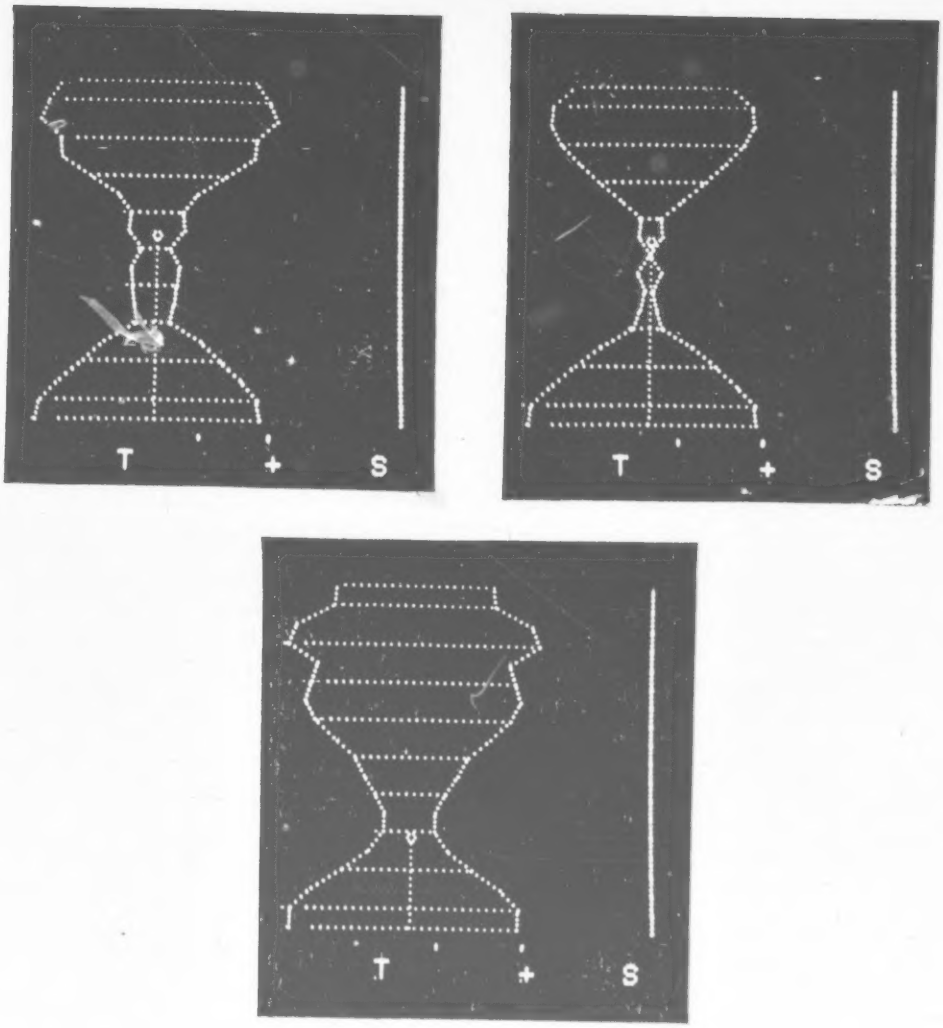


Figure 35. Vertical spectrum display for the vowel /i/, produced by three normally-hearing speakers. The horizontal lines indicate that the sound is voiced.

BEST COPY AVAILABLE

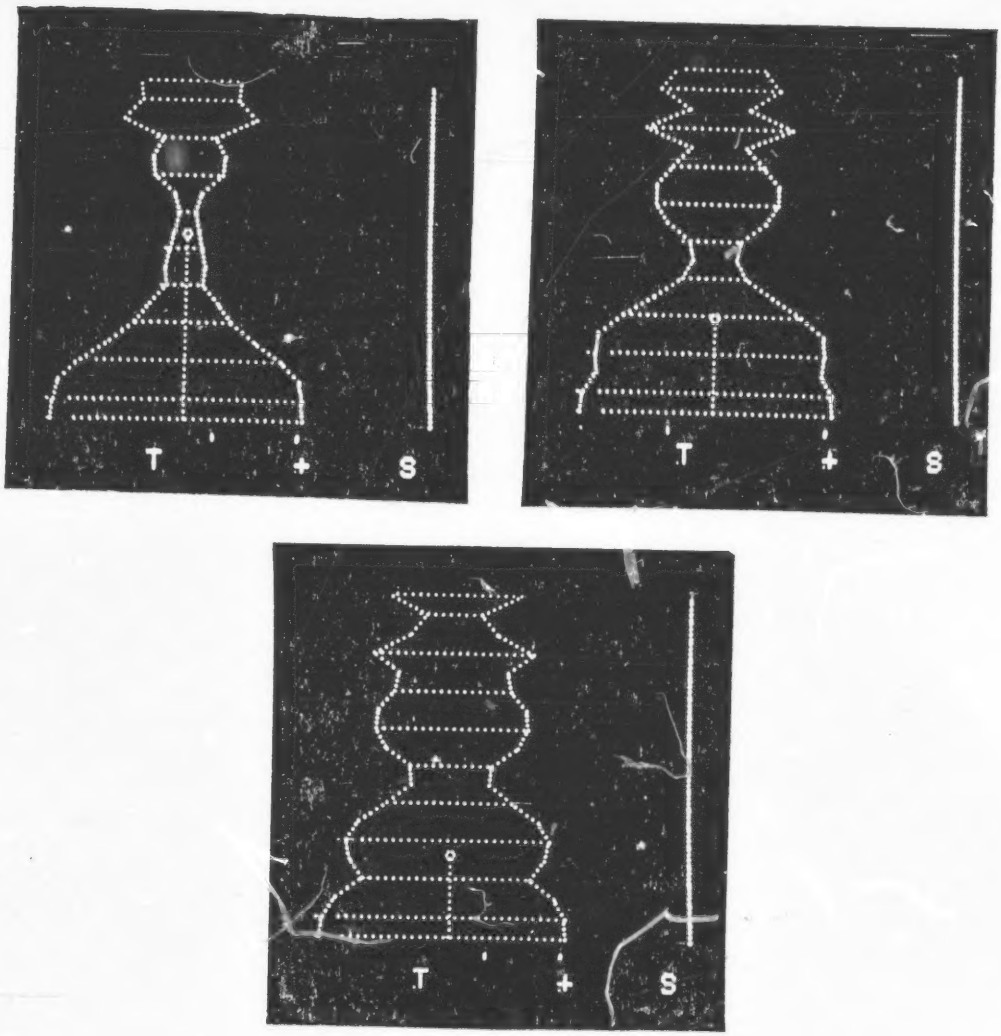


Figure 36. Vertical spectrum display for the vowel /u/, produced by three normally-hearing speakers. The horizontal lines indicate that the sound is voiced.

BEST COPY AVAILABLE

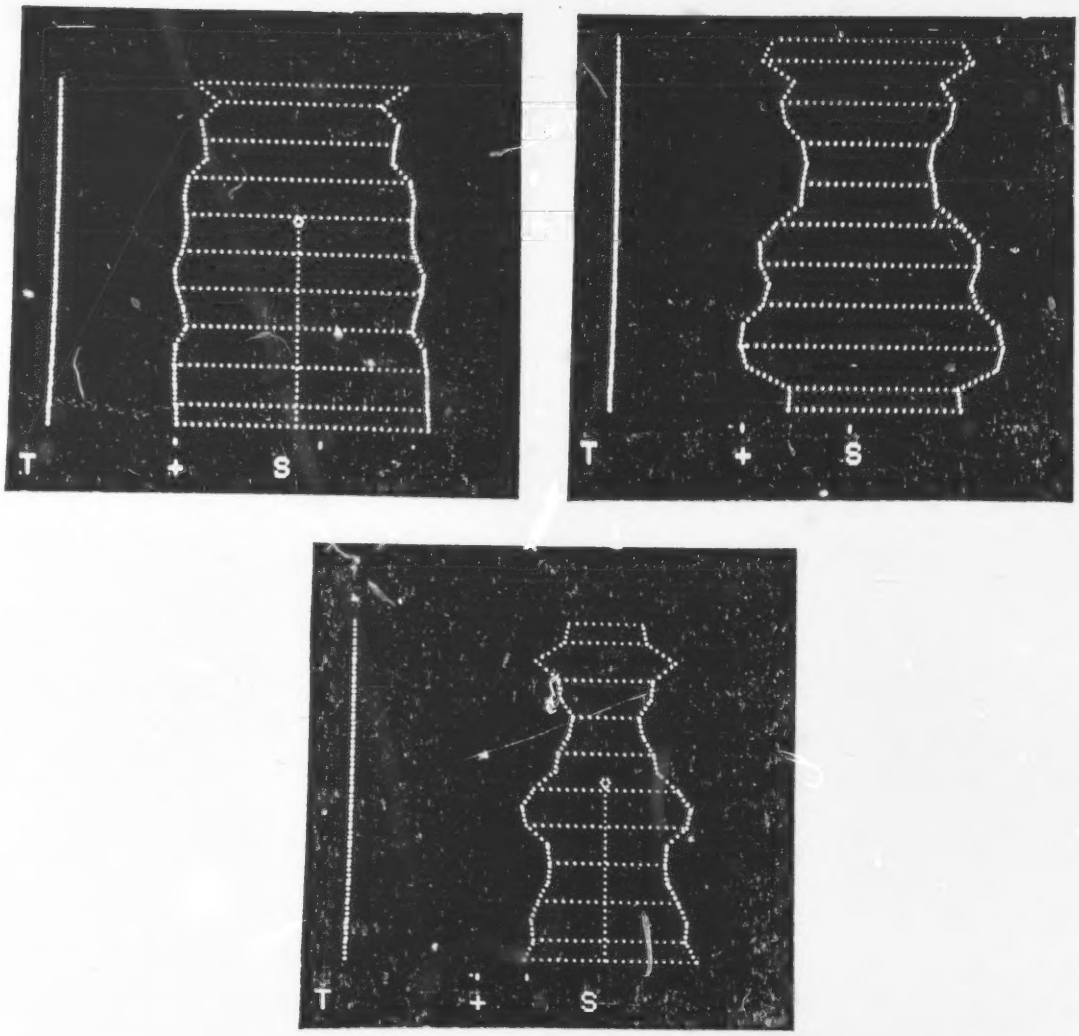


Figure 37. Examples of vertical display for the vowel /a/, produced improperly by deaf students.

BEST COPY AVAILABLE

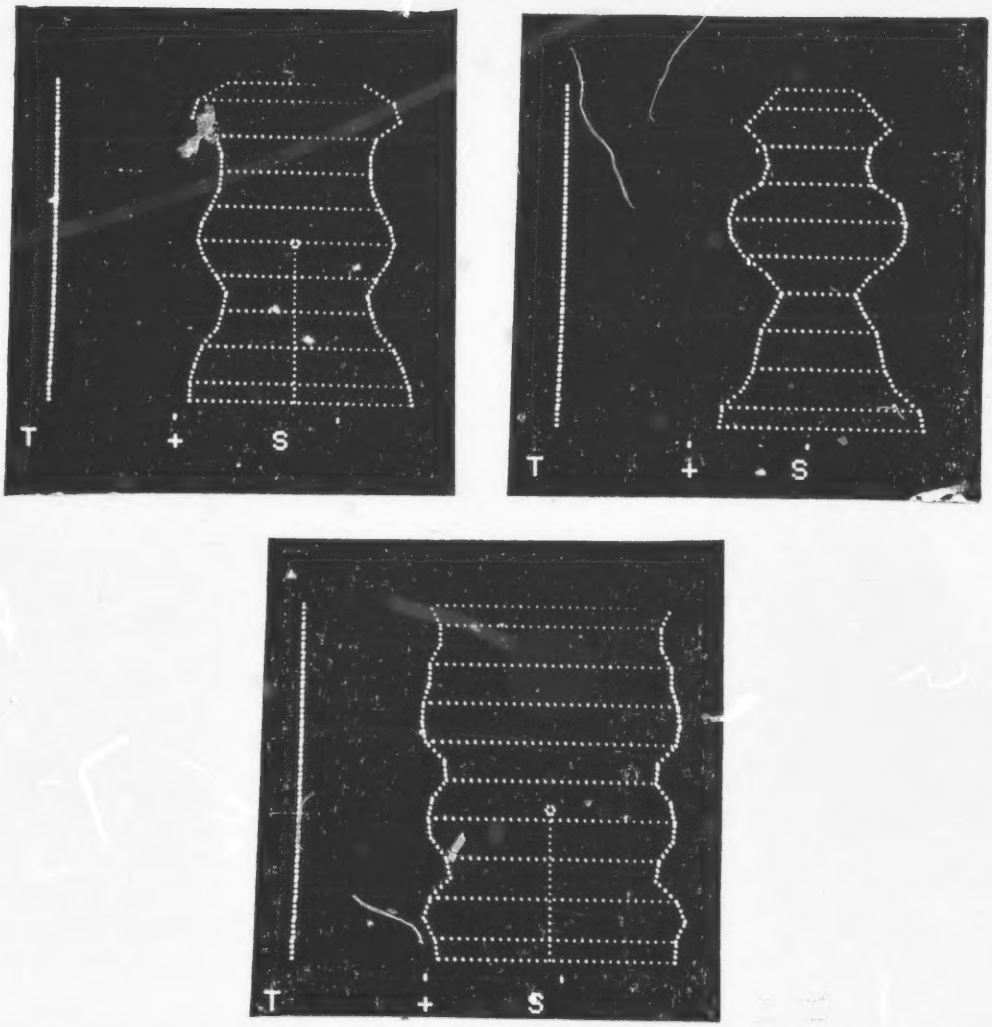


Figure 38. Examples of vertical display for the vowel /i/, produced improperly by deaf students.

in the pattern (Fig. 38b), or a maximum width in the lower part that is not at the bottom of the pattern (Fig. 38c).

The normal patterns for /u/ all have their widest point at the bottom. There is a broad section in the lower part of the pattern that narrows down gradually, yielding a minimum width near the middle. The lower half may sometimes include two bumps (Fig. 36c) and sometimes just one (Fig. 36a). There is a widening above the middle constriction, but the maximum width in the upper half is usually appreciably less than the width at the bottom. Typical errors in production (Fig. 39) lead to insufficient narrowing in the middle, and the widest point not at the bottom of the pattern.

These comments can be used as guides in working with students to produce approximations to these vowels produced in isolation. The normal procedure would be for the teacher to generate a display of the vowel, and point out the main distinctive attributes of the display. The student then attempts to produce a display with these attributes. If he is successful after some trials, the display of his own utterance can replace that of the teacher, and can be used by the student as a pattern to be matched on subsequent attempts.

The next logical step after work with isolated vowels is to shift to monosyllabic words produced in isolation. The procedure to be followed here is similar to that discussed above for fricative consonants in words. The word is produced, and the display during the vowel is observed, although this initial real-time impression is fleeting. The display for the word may then be replayed in slow motion and stopped during the vowel, thus permitting both student and teacher to examine the attributes of the pattern. A new utterance should then be made in an attempt to correct any inadequacies in the display.

BEST COPY AVAILABLE

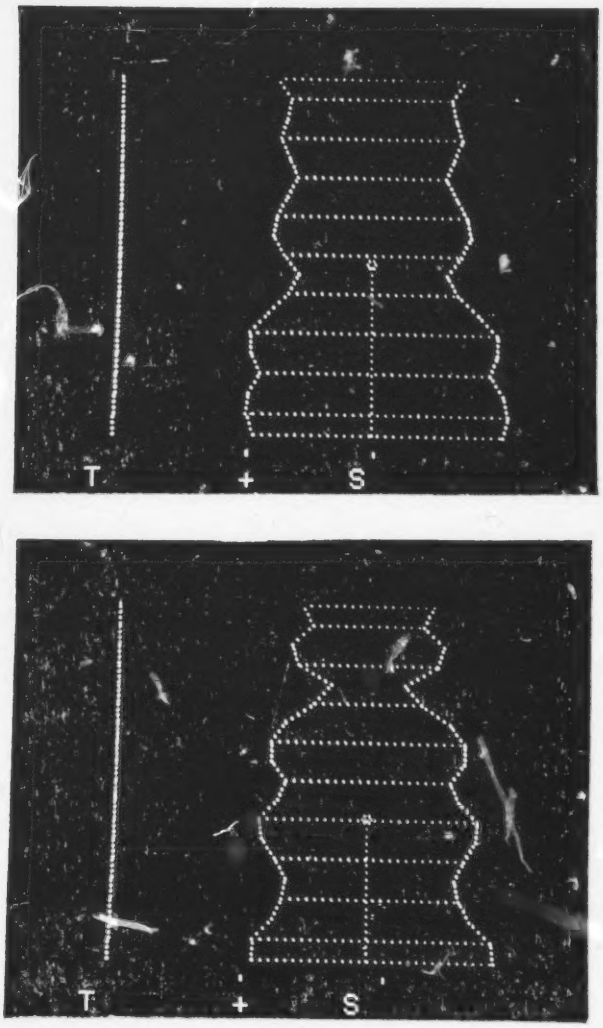


Figure 39. Examples of vertical display for the vowel /u/, produced improperly by deaf students.

Care must be exercised in following this procedure in working with vowel articulation in words. During the vowel portion of a consonant-vowel-consonant word, the articulation of the vowel may be influenced by the adjacent consonants, especially in regions of the vowel close to the consonants. Thus, observation of the vowel display should be made close to the middle of the vowel. For words with some types of consonants (e.g., /l, r, w, y/), the display does not show a well-defined boundary between vowel and consonant, with the result that the middle of the vowel may be difficult to find. When a word contains a nasal consonant (particularly in final position), the vowel may be nasalized, and the spectral patterns shown above for normal vowels will be modified. Training with vowel articulation in words should avoid words containing these classes of consonants.

The dual spectrum display can also be used to assist training in the production of diphthongs. When a diphthong like /ai/ (as in the word bite) is produced, for example, the display should change smoothly from /a/ toward /i/. Initial work with diphthongs should use diphthongs in isolation, first produced slowly and then more rapidly. Training can then be extended to diphthongs in isolated monosyllabic words.

4.5 Voice Quality

Study of the acoustic correlates of "voice quality" has not yet reached the point where deficiencies in voice quality can be diagnosed reliably solely on the basis of objective acoustic measurements. That is, there is no single acoustic measurement that shows good correlation with listener judgments of such qualities as breathiness or harshness. This difficulty may be due in part to the lack of agreement among listeners as to what constitutes a harsh or a breathy voice.

In spite of this lack of reliable acoustic correlates of breathy or harsh voice quality, a combination of listener judgments and objective measurements can be used with some success in the diagnosis of breathy voice quality and in the specification of training goals. Deaf students who are judged to have a very breathy voice quality also show a relatively high reading of the HL parameter for low vowels such as /a/, /æ/ and /ʌ/. Our data suggest that if the average HL in words containing these vowels is above 120, then the speech of that talker is judged by listeners to have a breathy voice quality. For these students, a display of the HL parameter for words with these vowels can be used to assess performance after some training.

A reasonable approach to the speech training of these students is to begin with the isolated vowel /a/. It should first be established that this vowel can be produced without excessive nasalization, following the guidelines indicated in Section 4.3. If there is excessive nasalization, it is appropriate to begin training to reduce nasalization before approaching the problem of breathy voice.

After it has been established that the nasalization is within normal limits, the HL reading for the isolated vowel /a/ should be determined. Let us assume that this reading is above 120 units for the HL display. The student should then be encourated to experiment with manipulation of his larynx in order to reduce the HL reading for this vowel. Several approaches are possible. The student might be instructed to let the air flow out of his lungs without effort and without forcing or tenseness in the muscles, letting the phonation occur on a natural exhalation produced by the elastic recoil of the lungs. He may also be encouraged to prepare for phonation with a closed glottis and with an initial pressure build-up behind the glottis. Phonation is then initiated

by a slight relaxation of this glottal closure, so that the glottal opening remains small. This kind of exercise should be carried out for a series of brief phonations, while the student is observing the result of his efforts on the HL display. He should aim for an HL reading well below 100.

Once the student has mastered the production of brief isolated vowels without excessive breathiness, he should shift to isolated monosyllabic words containing low vowels. Again, an HL criterion value well below 100 units should be set. The HL pattern for the teacher's utterance can be used as a guide for the student.

Some students tend to terminate an utterance with a breathy voice, probably as a consequence of premature abduction of the vocal cords before phonation ceases. This problem has been discussed in the section on timing and rhythm (Section 4.1), and examples of "loudness" displays for utterances with this problem were given. Work with students having this difficulty can also utilize the HL display, at least for low vowels. Training would utilize monosyllabic utterances with low vowels (such as /pa/), and students would be instructed to maintain a low HL reading throughout the utterance. Abnormal breathiness toward the end of the utterance would be manifested by an HL contour that curves up at the end of the vowel.

Only a few deaf students with breathy voice quality have worked with the HL display. Thus, it is not possible at this time to specify training procedures in greater detail. The teacher who has an opportunity to gain experience with this display should attempt to develop his own training procedures beyond those outlined briefly here.

5. ACKNOWLEDGMENTS

The contributions of several people to the development of the material in this manual deserve recognition. Charlene Long located many of the references that are cited in Section 2. Lawrence Prusak helped prepare much of the illustrative material that is included in Section 4. Helpful suggestions concerning training procedures came from several individuals, notably Lois Elliott, Patricia Archambault and members of the Clarke School teaching staff. Finally, both the patience of the Clarke School students and their consistent enthusiasm are acknowledged with pleasure; it is from the students that the real inspiration for such a project comes.

6. REFERENCES

Abrams, M. H., Goffard, S. J., Kryter, K. D., Miller, G. A., Sanford, J., & Sanford, F. H. Speech in noise: A study of factors determining its intelligibility. Harvard Univ., Psychoacoustics Laboratory, 1944. OSRD No. 4023, PB 19805.

Adams, M. E. The intelligibility of the speech of the deaf. American Annals of the Deaf, 1914, 59, 451-460.

Anderson, F. An experimental pitch indicator for training deaf scholars. Journal of the Acoustical Society of America, 1960, 32, 1065-1074.

Angelocci, A. A. Some observations on the speech of the deaf. Volta Review, 1962, 64, 403-405.

Angelocci, A. A., Kopp, G. A., & Holbrook, A. The vowel formants of deaf and normal hearing 11- to 14-year old boys. Journal of Speech and Hearing Disorders, 1964, 29, 156-170.

Bell, A. G. The mechanism of speech. New York: Funk and Wagnalls, 1916. 8th Ed.

Bodycomb, M. The speech of the deaf and the normal speaker. Volta Review, 1946, 48, 637-638.

Boone, D. R. Modification of the voices of deaf children. Volta Review, 1966, 68, 686-694.

Boothroyd, A. Concept and control of fundamental voice frequency in the deaf--An experiment using a visible pitch display. Paper presented at the International Congress of Education of the Deaf, Stockholm, Sweden, August, 1970.

Boothroyd, A. Acoustics of speech. In L. Connor (Ed.), Speech and the deaf child. Washington, D.C.: A. G. Bell, 1971, 3-44.

Borrild, K. Experience with the design and use of technical aids for the training of deaf and hard of hearing children. American Annals of the Deaf, 1968, 113, 168-177.

Brehm, F. E. Speech correction. American Annals of the Deaf, 1922, 67, 361-370.

Calvert, D. R. Some acoustic characteristics of the speech of profoundly deaf individuals. Ph.D. dissertation, Stanford Univ., Stanford, California, 1961.

Calvert, D. R. An approach to the study of deaf speech. Proceedings of the International Congress on Education of the Deaf, Washington, D.C., 1964, GPO, 242-245.

Calvert, D. R. Speech sound duration and the surd-sonant error. Volta Review, 1962, 64, 401-403.

Carhart, R. Development and conservation of speech. Chapter 14, in H. Davis & S. R. Silverman (Eds.), Hearing and deafness, Third Edition. Holt, Rinehart & Winston, 1970.

Carr, J. An investigation of the spontaneous speech sounds of five-year-old deaf-born children. Journal of Speech and Hearing Disorders, 1953, 18, 22-29.

Carr, J. Early speech development of deaf children. Proceedings of the International Congress on Education of the Deaf, Washington, D.C., 1964, GPO, 261-267.

Colton, R. H. & Cooker, H. S. Perceived nasality in the speech of the deaf. Journal of Speech and Hearing Research, 1968, 11, 553-559.

Curry, E. T. The pitch characteristics of the adolescent male voice. Speech Monographs, 1940, 7, 48-52.

Delattre, P. C. Two acoustic correlates of nasality studied by synthesis. Paper read before the Modern Language Association, Chicago, December 27, 1955.

Denes, P. On the statistics of spoken English. Journal of the Acoustical Society of America, 1963, 35, 892-904.

Dewey, G. Relative frequency of speech sounds. Cambridge, Mass.: Harvard University Press, 1923. Pp. 148

DiCarlo, L. M. The deaf. Englewood Cliffs, N. J.: Prentice-Hall, Inc., 1964.

Engelberg, M. Correction of falsetto voice in a deaf adult. Journal of Speech and Hearing Disorders, 1962, 27, 162-164.

Erber, N. P. Comment reported in R. E. Start (Ed.), Sensory capabilities of hearing-impaired children. Baltimore: University Press, 1974, 69-74.

Fairbanks, G. Recent experimental investigations of vocal pitch in speech. Journal of the Acoustical Society of America, 1940, 11, 457-466.

Fairbanks, G., Herbert, E. L., & Hammond, J. M. An acoustical study of vocal pitch in seven- and eight-year-old girls. Child Development, 1949, 20, 71-78.

Fairbanks, G., Wiley, J. H., & Lassman, F. M. An acoustical study of vocal pitch in seven- and eight-year-old boys. Child Development, 1949, 20, 63-69.

Fletcher, S. G. Theory and instrumentation of quantitative measurement of nasality. Cleft Palate Journal, 1970, 7, 601-609.

Fort, B. Spontaneous vocalizations of two-year-old deaf children as related to degree of hearing loss and other factors. Unpublished Manuscript, Univ. of Kansas, 1955.

Fry, D. B. Experiments on the perception of stress. Language and Speech, 1958, 1, 126-152.

Fujimura, O. Spectra of nasalized vowels. Quarterly Progress Report, Research Laboratory of Electronics, M.I.T., July 15, 1960, 214-218.

Guttman, N., Levitt, H., & Bellefleur, P. A. Articulatory training of the deaf using low-frequency surrogate fricatives. Journal of Speech and Hearing Research, 1970, 13, 19-29.

Hanley, J. D. An analysis of vocal frequency and duration characteristics of selected samples of speech from general American, eastern American and southern American dialect regions. Speech Monographs, 1951, 18, 78-93.

Heider, F., Heider, G. M., & Sykes, J. L. A study of the spontaneous vocalizations of fourteen deaf children. Volta Review, 1941, 43, 10-14.

Holbrook, A. & Crawford, G. H. Modification of speech behavior of the deaf (hypernasality). Paper presented to Conference of Executives of American Schools for the Deaf, St. Augustine, Florida, April, 1970.

Hollien, H. & Paul, P. A second evaluation of the speaking fundamental frequency characteristics of post-adolescent girls. Language and Speech, 1969, 12, 119-124.

Hollien, H. & Shipp, T. Speaking fundamental frequency and chronological age in males. Journal of Speech and Hearing Research, 1972, 15, 155-159.

Hood, R. B. Some physical concomitants of the perception of speech rhythm of the deaf. Ph.D. Thesis, Stanford, Univ., 1966.

Houde, R. A. Instantaneous visual feedback in speech training for the deaf. Paper presented at the 1973 Annual Convention of the American Speech and Hearing Association, Detroit, Michigan, October, 1973.

House, A. S. On vowel durations in English. Journal of the Acoustical Society of America, 1961, 33, 1174-1178.

Hudgins, C. V. A comparative study of the speech coordination of deaf and normal subjects. Journal of Genetic Psychology, 1934, 54, 3.

Hudgins, C. V. A study of respiration and speech. Volta Review, 1936, 38, 341-343, 347.

Hudgins, C. V. Voice production and breath control in the speech of the deaf. American Annals of the Deaf, 1937, 82, 338-363.

Hudgins, C. V. Speech breathing and speech intelligibility. Volta Review, 1946, 48, 642-644.

Hudgins, C. V. & Numbers, F. C. An investigation of intelligibility of speech of the deaf. General Psychology Monograph, 1942, 25, 289-292.

Huntington, D. A., Harris, K. S., Shankweiler, D., & Sholes, G. N. Some observations on monosyllable production by deaf speakers and dysarthric speakers. American Annals of the Deaf, 1968, 113, 134-146.

Irwin, D. O. Articulation of difficult speech sounds by two groups of retarded children. Cerebral Palsy Journal, 1966, 27, 3-7.

John, J. E. J. & Howarth, J. N. The effect of time distortions on the intelligibility of deaf children's speech. Language and Speech, 1965, 8, 127-134.

Joiner, E. Teaching of speech. American Annals of the Deaf, 1922, 67, 397-404.

Jones, C. "Deaf Voice." A description derived from a survey of the literature. Volta Review, 1967, 69, 507-508.

Klatt, D. H. Word verification in a speech understanding system. Invited paper, IEEE Symposium on Speech Recognition, Carnegie-Mellon Univ., April 15-15, 1974; to be published by Academic Press.

BEST COPY AVAILABLE

Levitt, H. Comment reported in R. E. Stark (Ed.), Sensory capabilities of hearing-impaired children. Baltimore: University Park Press, 1974, 73-76.

Licklider, J. C. R. & Pollack, I. Effects of differentiation, integration, and infinite peak clipping upon the intelligibility of speech. Journal of the Acoustical Society of America, 1948, 20, 42-51.

Lindblom, B. & Rapp, K. Some temporal regularities of spoken Swedish. Paper from the Institute of Linguistics, Univ. of Stockholm, Publication 21, September, 1973.

Lubker, J. F. & Moll, K. L. Simultaneous oral-nasal air flow measurements and cinefluorographic observations during speech production. Cleft Palate Journal, 1965, 2, 257-272.

Malvilya, M. P. Spontaneous vocalization and babbling in hearing impaired infants. Paper 16 presented at the seminar on Speech Communication Ability and Profound Deafness, International Congress on Education of the Deaf, Stockholm, 1970.

Mangan, K. R. Speech improvement through articulation testing. American Annals of the Deaf, 1961, 106, 391-396.

Martony, J. Studies on the speech of the deaf. Quarterly Progress and Status Report, Speech Transmission Laboratory, Royal Institute of Technology, January, 1966.

Martony, J. On the correction of the voice pitch level for severely hard of hearing subjects. American Annals of the Deaf, 1968, 113, 195-202.

Mason, M. K. & Bright, M. G. Tempo in rhythmic speech education. American Annals of the Deaf, 1937, 82, 385-401.

Menyuk, P. Hearing loss and language acquisition. Proceedings of the Multidisciplinary Symposium: Hearing Loss in Children, Children's Hospital Medical Center, 1972.

Miller, G. A. Decision units in the perception of speech. IRE Transactions on Information Theory, 1962, IT-8, 81-83.

Miller, M. A. Speech and voice patterns associated with hearing impairment. Audecibel, 1968, 17, 162-167.

Millin, J. P. Therapy for reduction of continuous phonation in the hard-of-hearing population. Journal of Speech and Hearing Disorders, 1971, 36, 496-498.

Murai, J. Speech development of a child suffering from a central language disorder. Studia Phonologica Journal, 1961, 58-69.

Mysak, E. E. Pitch and duration characteristics of older males. Journal of Speech and Hearing Research, 1959, 2, 46-59.

Neas, B. J. A study of the spontaneous babblings of three- and four-year-old deaf children. Unpublished Manuscript, Univ. of Kansas, 1953.

Nickerson, R. S., Kalikow, D. N., & Stevens, K. N. A computer-based system of speech-training aids for the deaf: A progress report. BBN Report No. 2901, September, 1974.

Nickerson, R. S. & Stevens, K. N. Teaching speech to the deaf: Can a computer help? IEEE Transactions on Audio and Electro-acoustics, 1974, AU-21, 445-455.

Nickerson, R. S., Stevens, K. N., Boothroyd, A., & Rollins, A. Some observations on timing in the speech of deaf and hearing speakers. BBN Report No. 2905, September, 1974.

Nober, E. H. Articulation of the deaf. Exceptional Children, 1967, 33, 611-621.

Parmenter, C. E. & Trevino, S. N. The length of the sounds of a Middle Westerner. American Speech, 1935, 10, 129-133.

Peterson, G. E. Influence of voice quality. Volta Review, 1946, 48, 640-641.

Philhour, C. W. An experimental study of the relationships between perception of vocal pitch in connected speech and certain measures of vocal frequency. Ph.D. Dissertation, Univ. of Iowa, 1948.

Phillips, N. D., Remillard, W., Bass, S., & Pronovost, W. Teaching of intonation to the deaf by visual pattern matching. American Annals of the Deaf, 1968, 113, 239-246.

Pickett, J. M. Sound patterns of speech: An introductory sketch. American Annals of the Deaf, 1968, 113, 120-126.

Pronovost, W. An experimental study of methods for determining natural and habitual pitch. Speech Monographs, 1942, 9, 111-123.

Quigley, L. F., Shiere, F. R., Webster, R. C., & Cobb, C. M. Measuring palatopharyngeal competence with the nasal anemometer. Cleft Palate Journal, 1964, 1, 304-313.

Rawlings, C. G. A comparative study of the movements of the breathing muscles in speech and in quiet breathing of deaf and normal subjects. Part I. American Annals of the Deaf, 1935, 80, 147-156. Part II. American Annals of the Deaf, 1936, 81, 136-150.

Rollins, A. M., Kalikow, D. N., & Nickerson, R. S. Operators Reference Manual and Software Documentation. BBN Report No. 2904, September, 1974.

Scuri, D. Restirazione e Fonazione nei Sordomuti. Rassegna di Educazione e Fonetica Biologica, 1935, 14, 82-113.

Shelton, R. L., Knox, A. W., Arndt, W. B., & Elbert, M. The relationship between nasality score values and oral and nasal sound pressure level. Journal of Speech and Hearing Research, 1967, 10, 542-548.

Smith, M. C. Residual hearing and speech production in deaf children. Communication Sciences Laboratory, Doctoral Program in Speech and Hearing Sciences, CUNY, April 15, 1973.

Sorenson, J. Fundamental frequency contours in the speech of deaf children. Term paper, M.I.T., Cambridge, Mass., 1974.

Stevens, K. N., Kalikow, D. N., & Willemain, T. R. The use of a miniature accelerometer for detecting glottal waveforms and nasality. BBN Report No. 2907, September, 1974. Submitted for publication to Journal of Speech and Hearing Research.

Stevens, K. N., Nickerson, R. S., Boothroyd, A., & Rollins, A. Assessment of nasality in the speech of deaf children. BBN Report No. 2902, September, 1974.

Stewart, R. B. The speech of children with high-frequency losses of hearing--a small phonetic study. Sound, 1969, 3, 40-43.

Stratton, W. D. Intonation feedback for the deaf through the tactile sense. M.I.T. Thesis, 1973.

Sykes, J. L. A study of the spontaneous vocalizations of young deaf children. Psychology Monograph. Studies in Psychology of the Deaf, 1940, 52, 104.

Templin, M. Certain language skills in children: Their development and interrelationships. Minneapolis: Univ. of Minnesota Press, 1957.

Thomas, I. B. The influence of first and second formants on the intelligibility of clipped speech. Journal of the Audio Engineering Society, 1968, 16, 182-185.

BEST COPY AVAILABLE

Voelker, C. H. An experimental study of the comparative rate of utterance of deaf and normal hearing speakers. American Annals of the Deaf, 1938, 83, 274-284.

Willemain, T. R. & Lee, F. F. Tactile pitch feedback for deaf speakers. Volta Review, 1971, 73, 541-554.